Call - EuroHPC AI Factory Antennas

HORIZON-EUROHPC-JU-2025-AIFA-01

Overview of this call[1]

Proposals are invited against the following Destinations and topic(s):

Topics	Type of Action	Budgets (EUR million)	Expected EU contribution per project (EUR million)[2]	Indicative number of projects expected to be funded
		2025		
Opening (indicative): 09 April 2025 Deadline (indicative): 09 July 2025				
HORIZON-EUROHPC-JU-2025-AIFA-01: EUROHPC AI FACTORY ANTENNAS	HORIZON-JU-RIA	70	5	7-8
Overall indicative budget		70		

General conditions relating to this call

The aim of AI Factories is to provide European startups as well as the industrial and the scientific community with enhanced access to AI optimised computing capabilities and underpinning services for the training and development of general-purpose, large-scale AI models, and for the development, testing and validation of emerging AI applications.

Call EUROHPC-2024-CEI-AI-01 and Call EUROHPC-2024-CEI-AI-02 were launched on 14 September 2024. The present Call for proposals aims to select 'AI Factory Antennas' that will be linked with an established AI factory and its AI optimised supercomputer. This call is launched in accordance with the EuroHPC Regulation[3], taking into account the EU Financial Regulation[4] and where relevant on the basis of Financial Rules of the EuroHPC JU[5].

[1] AI Factory Antennas

2 Nonetheless, this does not preclude submission and selection of a proposal requesting different amounts.

[3] Council Regulation (EU) 2021/1173 on establishing the European High Performance Joint Undertaking, as amended by Council Regulation 2024/1732 of 17 June 2024 amending Regulation (EU) 2021/1173 as regards a EuroHPC initiative for start-ups in order to boost European leadership in trustworthy artificial intelligence (https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202401732&qid=1741078538770)

[4] Regulation (EU, Euratom) 2024/2509 of the European Parliament and of the Council of 23 September 2024 on the financial rules applicable to the general budget of the Union (recast), OJ L, 2024/2509, 26.9.2024 (hereinafter referred to as 'FR') (Regulation - EU, Euratom - 2024/2509 - EN - EUR-Lex)

[5] Decision of the Governing Board of the EuroHPC JU No 3/2020 Approving the Financial Rules of the EuroHPC Joint Undertaking readopted by Decision of the Governing Board of the EuroHPC JU No 17/2021 approving the re-adoption of Governing Board Decisions

The aim of an AI Factory Antenna is to provide national AI ecosystems of a Participating State of the EuroHPC JU with relevant support services, algorithmic support, training activities and access to talent. The AI Factory Antenna will also have to ensure access to supercomputing resources from an established AI Factory.

This call for proposals will allow for more EuroHPC Participating States to have their own 'AI Factory Antenna' without having to invest in supercomputing resources required for the establishment of a fully-fledged AI Factory.

To be eligible under this call, proposals by EuroHPC Participating States for establishing an AI Factory Antenna must fulfil the following two conditions:

- **The Participating State has not yet been selected to host an AI Factory or be already partner of a selected AI Factory;**
- **The proposal must include a (pre- agreed) association with an established AI Factory.**

All selected AI Factory Antennas under this call will thus be in a position to provide AI-optimised computing capabilities and services for the large-scale training and development of general-purpose AI models, and other AI applications/ solutions to their national stakeholders and users of their countries. Where relevant, an AI Factory Antenna may under this call also acquire and provide (small scale) AI compute resources for the fine tuning, testing and validation of AI applications, complementing the AI-optimised computing resources provided by its linked AI Factory,

Proposals for establishing AI Factory Antennas are invited for the following topic(s):

HORIZON-EUROHPC-JU-2025-AIFA-01: EUROHPC AI FACTORY ANTENNAS

Specific conditions	
Expected EU contribution per AI Factory Antenna	The EuroHPC JU estimates that a maximum EU contribution of up to EUR 5 million for a duration of up to 3 years would allow these outcomes to be addressed appropriately. The duration should be aligned with the duration of the established AI Factory. .
Expected relationship between AI Factory and AI Factory Antenna	The established AI Factory will be expected to link up with the AI Factory Antenna and include it in all AI Factory activities. For these activities, it will not receive additional EU funds. The selected AI Factory Antenna and the AI Factory will be expected to sign a Memorandum of Understanding setting out the terms of their collaboration
Indicative budget	**The total indicative budget for the AI Factory Antenna call is up to EUR 70 million. The maximum amount of the EU contribution (up to 50 % of the total eligible costs) that may be allocated to an AI Factory Antenna is up to EUR 5 million and is subject to EU budget availability.**
Type of Action	HORIZON JU Research and Innovation Action

adopted under the framework of Regulation (EU) 2018/1488 and its updated Rules of Procedure in the view of Regulation (EU) 2021/1173.

Eligibility conditions	**The current Call is open to a single entity or a consortium of entities from one EuroHPC Participating State.** **Furthermore, entities established in countries with which association to Horizon Europe negotiations are being finalised and/or pending ratification, and which have committed to join the EuroHPC JU in the year when the call is launched, are eligible to participate. However, even if such entities are selected, the Grant Agreement can only be signed if the association with the Horizon Europe has started producing legal effects, and if the Governing Board of EuroHPC JU has confirmed membership of the new Participating State.** **Entities or consortia of entities that are already beneficiaries of an AI Factory grant and are fulfilling the conditions specified in Article 9 of the EuroHPC Regulation are ineligible to apply for this Call.** **Participation in more than one AI Factory Antenna is not permitted.**
Procedure	The granting authority will fund applications that meet all the evaluation criteria and subject to the ranking list approved by the EuroHPC Governing Board.
Legal and financial set-up of the Grant Agreements	Applicants to this call will need to submit the following documents : Applicants must provide a Letter(s) of Support proving the commitment of the Participating State to the AI Factory Antenna proposal and the commitment of this Participating State to cover the expenses of the proposal not covered by the EU contribution. Such commitment letter should be provided by an authorised representative of the Participating State. Applicants representing an AI Factory Antenna must provide a Letter of Intent co-signed with the consortium leader of an established AI Factory agreeing to their association, setting out how it will be linked with this established AI Factory. The AI Factory Antenna will be expected to substantiate how it can work in close cooperation and synergy with the established AI Factory extending the services it provides currently and in the future. This 'Letter of Intent' should also include an Annex providing a short overview of the main association conditions agreed between the respective parties and should include the financial conditions of such an association (if any). If the proposal is selected, this 'Letter of Intent' will be turned into a 'Memorandum of Understanding' (see below). In order to assess the applicants' eligibility, the following supporting documents are requested: - The legal entity identification form[6] duly completed and signed by the person authorised to enter into legally binding commitments on behalf of the applicant organisation(s) to be submitted in original; - AI Factory Antenna consortium: in addition to the supporting documents referring to their legal status, the consortium members will submit a signed declaration (modal mandate letter Annex XYZ) based

[6] http://ec.europa.eu/budget/contracts_grants/info_contracts/legal_entities/legal_entities_en.cfm

	on the model Consortium Agreement/Power of Attorney, appointing a consortium leader and giving a mandate to him. - Each applicant in an AI Factory Antenna consortium must fill-in and provide the duly signed Declaration of Honour
Other conditions	This action is an EU Synergy call. Grants and procurements can be linked with another grant funded from any other EU funding programme. The grants of this call will be managed as linked actions with the Calls EUROHPC-2024-CEI-AI-01 and EUROHPC-2024-CEI-AI-02

Expected Outcome:

The present Call for proposals is launched for the selection of an entity or a consortium of entities to establish an AI Factory Antenna. This call is launched in accordance with the EuroHPC Regulation, taking into account the EU Financial Regulation and where relevant on the basis of Financial Rules of the EuroHPC JU.

For the purposes of this call, an AI Factory Antenna is an entity, or a consortium of entities from a single EuroHPC Participating State.

The AI Factory Antennas will strengthen the network of EuroHPC AI Factories. The AI Factory Antennas will provide relevant support services, algorithmic support, training activities, access to talent and other support services to the national AI ecosystem(s) while also ensuring remote computing access to AI-optimised supercomputing capacity of the linked AI Factory. These services will also contribute to the mission and objectives of an 'established AI Factory' to which the AI Factory Antenna is associated.

An 'established AI Factory' is an AI Factory which has been selected via Calls EUROHPC-2024-CEI-AI-01 and EUROHPC-2024-CEI-AI-02. For the purposes of this call, an 'established AI Factory' may also be referred to as an 'AI Factory hosting entity'

The call is open to entities or consortia of entities fulfilling cumulatively the following activities as defined in Article 4 (h) (iv), (v), (vi), (vii), (viii) and conditions as defined in Article 9 (5a) of the EuroHPC Regulation:

a) The applicant entity or entities shall represent one single EuroHPC Participating State that has agreed, with a Letter of Intent, to contribute to participating in an established 'AI Factory'.

b) The applicant entity that represents the 'AI Factory Antenna' and the coordinator of the established AI Factory shall enter into an agreement which, if successful, will form the basis of Memorandum of Understand to this effect.

c) The applicant entity that represents the 'AI Factory Antenna' has to be registered as a legal entity in one of EuroHPC's Participating States.

d) The applicant entity or entities that are identified in the application as being part of the 'AI Factory Antenna' must have a legal personality on the date of the deadline for submission of applications and must be able to demonstrate their existence as a legal person. In case the application is submitted by several different legal entities from the same Participating State, this criterion (c) applies to all entities.

e) Applications should include the provision of appropriate supporting documentation proving the commitment of the Participating State and to cover the share of the total cost of the participation in the established AI Factory that is not covered by the Union contribution or by the hosting

member state where the established AI factory is located as set out in Article 5 of the Regulation or any other Union contribution as set out in Article 6 of the Regulation. Such commitment documentation should be provided by an authorised representative of the Participating State. A commitment letter template will be provided for applicants (to be added later).

~~f)~~ In the case of an AI Factory Antenna 'consortium', the agreement shall take the form of a partnership between the legal entities from the Participating State, of which the consortium leader will take the lead and act as coordinator of the consortium. The coordinator will act as an intermediary for all communications between the established AI Factory, the EuroHPC JU and the AI factory Antenna. However, if selected, the partners of AI factory Antenna consortium as well as the Hosting Entity of the AI Factory are jointly responsible for implementing the action(s) resulting from the awarded grant agreement. To implement the action(s) properly, they must make appropriate internal arrangements. The AI Factory Antenna must be given power of attorney by the applicant entities to represent it to sign and administrate the grant agreement (consortium leader).

g) The AI Factory Antenna shall assume full liability towards the EuroHPC JU for the implementation of the AI Factory Antenna services as a whole, including financial and operational liability.

Scope:

Applicants should provide:

a) A description of the national AI ecosystem(s) their AI Factories Antenna aims to target

b) A description of the features their AI Factories Antenna will have in terms of activities and service offering.

c) A description of how the AI Factory Antenna will cooperate and synergise with the established AI Factory it will be linked with, extending/enhancing/complementing the services and activities it provides currently or in the future.

In detail the Applicants should structure their proposal as follows:

a. A general description of the concept of the proposal and the needs for an AI Factory Antenna

1) *Concept of the AI Factory Antenna*

   a) *Vision, Rationale and Objectives* of the proposed AI Factory Antenna.

   b) *A roadmap* for developing the national AI ecosystem(s) and how it would be served, justifying the need for setting up the AI Factory Antenna.

2) *Targeted key Industrial sectors and Applications and targeted Stakeholders and their needs*:

   a) Description of the *key industrial/application sectors* as well as of the key obstacles to overcome to further develop the AI innovation ecosystem in these sectors.

   b) Presentation of a convincing plan for attracting key AI stakeholders from these sectors.

   c) Description of any plans the Applicants to include *internal or external cloud solutions* to bridge the needs towards an end-to-end computing continuum.

3) *Links to a national AI strategy, and national data and access policies to computing and data*:

a) Description of how the AI Factory Antenna proposal is linked to the national AI Strategy / Strategies or equivalent[1] of the Applicant(s).

b) Description of how the AI Factory Antenna is linked to a current *National Data Policy* of the hosting entity or the hosting consortium, enabling access to large datasets. If this does not exist, description of a plan to make available large data sets to the AI Factory Antenna ecosystem.

c) Where relevant, description of an *AI user-friendly access policy of the AI Factory Antenna* to the national share of computing time of the associated EuroHPC supercomputer and how it will contribute to the development of the national AI Ecosystem.

4) Unless otherwise agreed with the established AI Factory, Applicants must *outline an indicative financial contribution plan,* specifying how the AI Factory Antenna will support the established AI Factory in terms of, for example, in kind or in cash contribution [for example, participation in the capital expenditure (Capex) or operational expenditure (Opex) of the AI optimised supercomputer of the established AI Factory] for access to the AI-optimised supercomputing resources and to any of the respective services of the established AI Factory. A preliminary agreement between the AI Factory Antenna and the established AI Factory may be documented in the Letter of Intent. In the Annex of this 'Letter of Intent', a short overview of the main association conditions agreed between the respective parties should be described as well as the financial conditions of this association (if any). If the proposal is selected, the 'Letter of Intent' will be formalised into a Memorandum of Understanding (MoU) upon selection.

5) Overall plan for networking the AI Factory Antenna with the other EuroHPC AI Factories and Antennas.

6) *Overall plan for linking the AI Factory Antenna to a national strategy for startups/SMEs:* description of the plans the Applicants have for linking the AI Factory Antenna ecosystem with relevant national/regional investment measures targeted at startups and SMEs.

b. A detailed description of the AI Factory Antenna data facilities and services and its networking with other AI Factories and AI Factories Antennas:

1) *AI Factory Antenna tools and services*

- *Overview of the user support services*: This includes: (i) Description of the range of services that the AI Factory Antenna will provide to the AI ecosystem (e.g., guidance for using the HPC environment, adapting the computational tasks associated to the training and fine-tuning of the AI models and related inference activities to the HPC environment, etc.). (ii) Description of a plan for servicing private and public national users as well as users from other EuroHPC Participating States. (iii) Description of the foreseen professional user support plan, describing the range of user support activities (i.e., how the AI Factory Antenna plans to engage with and serve the broader AI community – from startups, SMEs and large industry to academia and research institutions – and how will these professional services be provided). (iv) Description of the resources required to support the established AI Factory in providing a well-functioning user support service.

- *Computing, Software and application development environments*: description of any (small scale) AI computing resources available or to be acquired by and become available under the AI Factory antenna, as well as the software environment the AI Factory Antenna will

[1] In the absence of a formal national AI strategy, the Applicants will need to describe the strategic national (or Consortium) character of their AI Factory Antenna approach.

support the AI Factory in delivering, including ready-to-use set of AI-oriented tools containerized workloads and workflows, etc.

2) *Data facilities, access to data, confidentiality and integrity of data*

   - *Data facilities:* Description of the data repositories and data assets that the AI Factory Antenna plans to make available to the AI ecosystem.

   - *Access to Common European Data Spaces*, including preliminary agreements on the principles of an access and use, establishing relevant data repositories (e.g., Hugging Face).

   - *Plans for establishing secure and trusted environments,* for guaranteeing the confidentiality and integrity of sensitive data and for ensuring the integrity of computational processes.

3) *Trustworthy AI*: description of the plans the Applicants have for developing of robust guidelines and standards for AI algorithmic development aligned with the principles and requirements of the AI Act.

4) *AI Factory Antenna Hub facilities (where relevant)*

   - *co-working space facilities:* description of the plans the Applicants have for making available co-working space physical facilities, possibly complemented also by virtual working spaces.

   - *hosting facilities for AI students*: and description of the Applicants Plans for making available a physical campus hosting AI students located nearby or associated to the established AI Factory.

5) *AI Factory Antenna training facilities*

   - *Skills plan*: Description of the AI Factory Antenna Skills Plan outlining the skills needed for the targeted AI stakeholders, including a description of a diverse range of training courses, complementary training facilities and activities and timelines tailored to the varying needs of potential users.

   - *Access to human capital*: in house and external direct access to the necessary human capital and talent to provide the necessary education/training activities planned. This includes plans for collaboration and engagement with universities to train and equip students at all levels with the necessary in-demand AI skills.

6) Detailed plans for networking the AI Factory Antenna with existing European and national initiatives and with other EuroHPC AI Factories and AI Factory Antennas.

   - *Networking with other existing European and national AI & HPC initiatives:* Detailed plans for linking the AI Factory Antenna with European and national AI and HPC initiatives such as TEFs, EDIH, National HPC Competence Centres, ALT-EDIC, or others, and to engage with them while avoiding duplication of efforts.

   - *Networking with other AI Factories and AI Factories Antennas*: Detailed plans for linking the AI Factory Antenna with other established AI Factories and AI Factory Antennas once they become operational in order to network, exchange best practice, share experiences, and avoid duplication of efforts.

c. A description of the AI Factory Antenna Implementation Plan:

1) *Implementation plan and risk management*: Applicants should provide an indicative implementation plan, an organisational structure and roles for the management of the AI Factory Antenna, and a project timeline with phases for the establishment of the AI Factory Antenna. Applicants should also include a risk management approach by identifying potential risks and mitigation strategies. Applicants should also indicate a timeline for the signature of a Memorandum of Understanding with the established AI Factory to take place if selected. This Memorandum of Understanding should set out agreed activities and milestones to deliver these activities.

2) Key performance indicators (KPIs): Applicants should provide a description of a set of KPIs and metrics that the Applicant(s) will use to measure the contributions to the success of the activities of the AI Factory Antenna and associated AI ecosystem. (see European Commission Concept Paper found in the Annex 1 of this Call for Proposals for more information)

3) *Budget estimate of the proposal*: Applicants should provide an estimated budget the establishment of the AI Factory Antenna, including development, implementation and expected operational costs.

d. A comprehensive description of the expected Impacts of the AI Factory Antenna:

Applicants should describe the pathways to achieve the expected outcomes and expected impacts and the measures they will take for maximising these expected outcomes and impacts.

Evaluation and Selection Criteria

This Call for Proposals will be evaluated by a peer review process against the evaluation criteria detailed below:

a. **Vision, plans and capability of the AI Factory Antenna to address the challenges of the Artificial Intelligence start-up ecosystem, and research and innovation ecosystem and the Artificial Intelligence user community and providing a supportive centralised or distributed Artificial Intelligence-oriented supercomputing service**

- Contribution, clarity and pertinence of an AI Factory Antenna being linked to an established AI Factory, in terms of vision, rationale, objectives, development roadmap, targeted key industry sectors and stakeholders, internal or external cloud solutions planned to bridge the needs towards an end-to-end computing continuum and networking with other initiatives.

- Contribution, clarity and pertinence of an AI Factory Antenna being linked to an established AI Factory data facility, access to data, confidentiality and integrity of data.

- Pertinence of the links of the AI Factory Antenna to the respective national AI Strategy, national data and access policies to computing and data, and to a national strategy for investing in startups/SMEs.

- Quality and efficiency of the Implementation Roadmap, including its deliverables and milestones, the risk management approach and the Key performance Indicators.

- Clarity and pertinence of the plans to invest in physical and virtual infrastructure, including in (small scale) AI computing resources, required for the AI Factory Antenna.

- Soundness of the AI Factory Antenna's budget.

- Credibility of the pathways to achieve the expected outcomes and expected impacts.

- Suitability and quality of the measures to maximise expected outcomes and impacts.

b. Quality and pertinence of experience and know-how available from the applicant entity that would provide Artificial Intelligence-oriented supercomputing service environment

- Quality and pertinence of experience and know-how available from the applicant entity that will support the established Artificial Intelligence-oriented supercomputing service environment.
- Quality and pertinence of the AI Factory Antenna user support services, including the quality and efficiency of the plan for offering professional services.
- Quality and pertinence of the AI Factory Antenna tools and software and application development environments.

c. Plans for interaction and cooperation of the AI Factory Antenna with the established AI Factory, other established AI Factories, and other AI Factory Antennas with EuroHPC Competence Centres and EuroHPC Centres of Excellence, and with relevant Artificial Intelligence activities such as the hubs of Artificial Intelligence start-ups, the Artificial Intelligence and data ecosystems, the Artificial Intelligence Testing and Experimentation Facilities, the European central Artificial Intelligence platform, the Artificial Intelligence-oriented Digital Innovation Hubs and other related initiatives.

- Quality and pertinence of the proposed AI Factory Antenna.
- Clarity and pertinence of the networking activities of the AI Factory Antenna with established European and national initiatives and with other EuroHPC AI Factories.
- Soundness of the plans of the AI Factory Antenna for developing Trustworthy AI.
- Clarity on how the activities of the AI Factory Antenna are complimentary with the established AI Factory and not overlapping with established or future National Competence Centres (NCCs). As a reminder, the NCCs should establish and maintain a network of national HPC users, promote HPC use and uptake in the private and public sector and reach out to new potential users. This includes awareness raising and outreach activities to communicate the benefits of HPC to potential users with a specific focus on SMEs. The NCCs should foster the development of the necessary expertise, especially for HPC applications, of the local communities and relevant national stakeholders in collaboration with other NCCs and European initiatives. Furthermore, the role of NCCs is limited to providing HPC expertise, training, advisory and consultancy services rather than engaging directly in operational activities of clients and stakeholders.

d. Existing capabilities and future plans of the AI Factory Antenna to contribute to the development of the talent pool

- Pertinence and effectiveness of existing capabilities and future plans of the hosting entity to contribute to the development of the talent pool.
- Quality and pertinence of structured training facilities and training programmes highlighting relevant courses, activities, and learning pathways tailored to meet the diverse needs of potential users.
- Quality and pertinence of strategy to foster collaboration and engagement with universities, research centres and other training providers to train and equip students at all levels with the necessary in-demand AI skills.

Applicants should also refer for guidance to the European Commission Concept Paper found in the Annex 1 of this Call for Proposals.

ANNEX 1: "AI Factories" Concept Paper

Version 4.0, 25 July 2024

This concept paper addresses the EuroHPC Governing Board Members. It defines the way to implement the AI Factories[7]. It describes how the EuroHPC JU and Member States and consortia are to establish AI Factories and outlines their key features and activities. These will be reflected in the EuroHPC Call for Expression of Interest to host AI Factories.

Section 1 of this concept paper provides a description of what is an AI Factory. Thereafter a set of eligibility conditions for Member States to implement AI Factories are presented in Section 2. Section 3 provides a summary of the technical specifications that are expected to be addressed in Member States proposals on AI Factories. The Appendix I to this paper provides an overview of the different implementation modes to establish AI Factories across the EU through the EuroHPC JU.

1. What are AI Factories?

The Commission launched the AI Innovation Package in January 2024 to support European startups, and SMEs in the development of trustworthy AI. The AI Package proposed a limited number of targeted amendments to the EuroHPC JU Regulation for implementing the AI Factories around the EuroHPC supercomputers, which were largely endorsed by the Competitiveness Council on May 23, 2024.

The amended EuroHPC Regulation, so called the "AI Factories Act", expanded its objectives to include the development and operation of 'AI Factories'. AI Factories are entities which provide an AI supercomputing service infrastructure and will build open AI ecosystems formed around EuroHPC supercomputing facilities (hosting entities[8]). The activities covered by AI Factories will be open to public and private users, and with privileged access conditions for startups and small and medium-sized enterprises (SMEs). The amended regulation brings together the necessary resources around these supercomputers – namely computing power, data, and talent, to offer a wide and exhaustive range of services to public and private users, AI startups and SMEs, AI companies and researchers needed for the development of European general purpose AI models and other emerging AI applications or data driven applications, as well as subsequent targeted inferencing activities.

[7] According to the AI Factories Act (Council Regulation (EU) 2024/1732 of 17 June 2024 amending Regulation (EU) 2021/1173 as regards a EUROHPC initiative for start-ups in order to boost European leadership in trustworthy artificial intelligence), an AI Factory is a centralised or distributed entity providing an Artificial Intelligence supercomputing service infrastructure which is composed of: 1) an Artificial Intelligence-optimised supercomputer or Artificial Intelligence partition of supercomputer, 2) an associated data centre, dedicated access and artificial intelligence-oriented supercomputing services and attracting and pooling talent to provide the competences required in using the supercomputers for Artificial Intelligence. AI Factories should include the following features:

i. Acquiring, upgrading, and operating AI-optimised supercomputers to enable fast machine learning and training of large General Purpose AI (GPAI) models;
ii. Facilitating access to the AI dedicated supercomputers, contributing to the widening of the use of AI to a large number of public and private users, including startups and SMEs;
iii. Offering a one-stop shop for startups and innovators, supporting the AI startup and research ecosystem in algorithmic development, testing evaluation and validation of large-scale AI models, providing supercomputer-friendly programming facilities and other AI enabling services;
iv. Enabling the development of a variety of emerging AI applications based on GPAI models;
v. Attracting, pooling, and training talent to develop their competences and skills in using the EuroHPC supercomputers for AI.

[8] 'hosting entity' refers to a legal entity which includes facilities to host and operate a EuroHPC supercomputer and which is established in a Participating State that is a Member State.

AI Factories in each Member State or a hosting consortium of Participating States will be connected to those in other Member States and to other relevant AI initiatives, such as Testing and Experimentation Facilities, Digital Innovation Hubs, EDICs, etc., thus creating a closely interconnected AI ecosystem across the whole Europe.

The different elements of an AI Factory should not be seen in isolation but rather aligned and mutually reinforce each other. **The AI Factories** should cover **two main components namely i) the AI optimised Supercomputer and ii) the associated "AI Factories" activities and services**.

It is expected that a number of AI Factories will be established in a few Member States or consortia of Participating States around existing, upgraded or new AI optimised supercomputers. These AI Factories will serve the European and national AI communities.

The **AI Factories will be serving public and private users from all the EuroHPC Participating States**, including those which are not eligible or do not wish to host an AI Factory. Such users may be granted access to the share of EU's access time and necessary services provided by any of the EuroHPC AI Factories.

In order to serve users from Participating States, which do not host an AI Factory, the EuroHPC JU will act as first entry point. The JU will then dispatch the request to the appropriate AI Factory/Factories based on a number of selection criteria. These criteria as well as the access policy concerning the EU access time will be defined and agreed in due time by the EuroHPC Governing Board.

AI startups, which are supported through the EIC Acceleration Challenge of Horizon Europe, will be given a priority access to the AI optimised supercomputers and services offered by an AI Factory.

The EuroHPC Participating States, which do not host an AI Factory, can collaborate with one or more AI Factories through a strategic agreement with a hosting entity, similar to many of the current EuroHPC systems.

The provision of services by the AI Factories should be without prejudice to the EU **state aid rules**. The European Commission will provide guidelines in due time on this matter. In principle, provision of (free) services to startups and SMEs should be covered by the General Block Exemption Regulation[9]. On the other hand, provision of services to big industry should be fee-based.

1) AI Factories - Key Features to consider from a national perspective

The following section outlines a set of key policy features and technical activities that a Member State or a consortium of Participating States should undertake to support the development of an AI Factory that is to be co-funded by the EuroHPC JU. These are further summarised in Appendix II and will be further expanded in the relevant Calls for Expression of Interest.

Investing in AI optimised supercomputers

AI Factories should be developed around AI optimised supercomputers to address and serve the needs of national users, their AI ecosystem and potential AI European and national AI stakeholders and serve the needs of their targeted AI ecosystem. There are three possibilities that a Member State or a

[9] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014R0651.

Consortium of Participating States and the corresponding hosting entity can consider here – these are presented in detail in the Appendix I.

Creating a national AI Ecosystem

Hosting entities should define and justify their needs and design choices in the context of their targeted AI usages and national/local ecosystems. Hosting entities should therefore present a comprehensive assessment of the users and AI Ecosystems they would like to serve and enhance through the AI Factory, ensuring a minimum critical mass justifying the need for an AI Factory. The assessment should include at least the following aspects:

i. **National AI Strategy**

To what extent the establishment and deployment of an AI Factory is linked and contributes to the implementation of the national AI strategy of the hosting country/countries of the hosting consortium.

ii. **National Data Policies and access strategies to the AI optimised supercomputers**

Applicants would need to describe the current National Data Policies in place (including possible access to data spaces that are available through their participation to EU initiatives such as EDICs) for enabling the access to large datasets, as well as the availability of knowledge corpus. In cases where such National Data Policy does not exist, applicants would need to provide a plan of how they will make available large data sets to the AI Factory ecosystem.

In both the above cases, Applicants should describe how they will implement policies facilitating the access to open / FAIR[10] data and proprietary data (including if necessary different fee schemes depending on the use of data for AI training/fine-tuning/inference).

iii. **National Access Policy to AI Community**

To ensure a cohesive HPC for AI approach and foster the national and local ecosystem, it is expected that Applicants would put in place an AI user-friendly access policy to the national share of computing time of the EuroHPC supercomputer and describe how it will contribute to the development of the national AI Ecosystem.

iv. Stakeholders

To build a thriving AI ecosystem, Applicants should clearly identify and be capable of attracting key stakeholders which can contribute to their success of their AI ecosystem. These should include:

a. AI Companies/AI Developers/AI Startups and SMEs.

b. AI Technology solution providers.

c. Potential Data providers which can supply high-quality data for AI training and analysis.

d. AI Users that will benefit from AI Factories generated AI-driven applications and solutions.

e. AI communities, including academia and students.

[10] Findable, accessible, interoperable, and reusable

f. Private investors / incubators.

v. AI Ecosystem needs and challenges

Applicants should identify the needs and challenges of the AI ecosystem they intend to serve. Each AI Factory should preferably focus on selected applications/domains that are aligned with the strategic vision and strategic specialisation areas of the hosting country and/or the consortium of Participating States. They should identify the key barriers and obstacles that may hinder the creation of a thriving AI ecosystem, and the extent to which the deployment of the AI Factory can overcome these obstacles to create an AI ecosystem that harnesses the full potential of AI for the benefit of the relevant stakeholders.

Applicants may include internal or external cloud solutions to bridge the needs towards an end-to-end computing continuum spanning model development, training, fine-tuning, and inference.

vi. Strategy for AI startups/SMEs

To foster a thriving AI ecosystem, a proactive startups / SMEs policy at a national/regional level plays a vital role in fostering and attracting investment in the AI sector. By facilitating access to capital to startups/SMEs and/or implementing targeted tax incentives, governments can encourage investment and support startups/SMEs to ensure the success and growth of businesses. Hosting entities are encouraged to link the AI Factory ecosystem with relevant national/regional investment measures targeted at startups and SMEs.

vii. AI Factories – KPIs

Applicants should propose key performance indicators (KPIs) and metrics to measure the contributions to the success of their AI Factory and associated AI ecosystem, such as (but not limited to):

- Number of private AI users served annually, notably start-ups and SMEs;
- Number of public AI users served annually.
- Number of participants in the AI Factory ecosystem, including European ones, served.
- Number and quality of services provided by the AI Factory
- Number of AI training sessions provided.
- Number/quality/size of GenAI open models released.
- Number/volume of available quality databases annually.
- Number of AI science applications served/released.
- Number of industrial/SME/startup applications served/released.
- Number of AI dedicated researchers in the AI Factory.
- Number of students participating in AI Factory activities.
- Usage of the AI optimised supercomputer.

- Degree of oversubscription to the AI access calls.

Applicants may propose other relevant KPIs.

2) Overview of the Technical Specifications / Activities of AI Factories

This section provides a succinct overview of the main technical aspects that are expected to be included in the forthcoming AI Factory Calls for Expression of Interest.

a. Compute

AI Factories should deliver a minimum computing capacity to address the needs of users and their AI ecosystem, including potential AI European model developers and serve the needs of their targeted AI ecosystem.

Their targeted compute requirement should be ideally justified through the use AI/HPC benchmarks. These may include, e.g. (indicative)**:**

- **HPL-MxP benchmark:** The high-performance Linpack mixed precision benchmark seeks to address the convergence of HPC and AI workloads.
- **MLPerf Training HPC benchmark:** Benchmark, targeted at supercomputers, measuring the performance of training machine learning models for scientific applications and data**.** Minimum time-to-solution (e.g., training a 10 billion parameter language model in 45 days).

Applicants may propose further benchmarks, including inference related benchmarks where appropriate.

b. Storage

AI Factories must ensure enough storage capacity to handle large and numerous databases, as well as providing the necessary flexibility to increase their capacity according to the evolution of needs of users. **The storage should be collocated with the supercomputer or connected through a high speed (terabit) connection to maximize data throughput and minimise latency.**

- **High-capacity storage:** Adequate storage capacity to manage vast datasets**.**
- **High-speed storage**: Availability of fast storage to ensure rapid data access and transfer.

Applicants are expected to propose I/O[11] benchmarks to test the performance of proposed storage systems.

To strike a balance between capacity and speed, a tiering storage approach that combines different technologies, from fast disks to tapes, may be considered.

c. Data

[11] Input/output operations.

The availability and accessibility to large data repositories with high quality curated data is fundamental for the AI community to flourish. AI Factories must guarantee high-speed connectivity and unrestrained access to European Data Spaces and relevant data repositories.

- **Data facility:** Co-located or very high-speed connection to (at least) one associated data facility linked to the supercomputer. Data centres to host large volumes of data necessary for AI Factories and associated data facilities must be operational within 12 months of being selected to host an AI Factory.
- **Access to Common European Data Spaces[12]:** Hosting entities should clearly identify interaction with and access to which Common European Data Spaces they wish to interact and have access to, provided that these correspond to their targeted / selected applications / domains that are aligned with the strategic vision and strategic specialisation areas of the hosting country / hosting Consortium. Hosting Entities should also describe the principles of an eventual access to and use of agreement with such Common European Data Spaces. Complementary and relevant data repositories (e.g., Hugging Face) should also be considered, as well as readiness to integrate into the future EuroHPC Federation Platform, which will be federating EuroHPC JU supercomputers and European HPC resources.
- **Security:** AI Factories should guarantee the confidentiality and integrity of sensitive data and ensure the integrity of computational processes. Users of computing capacity could for example be authenticated using the EU eID Wallet, once available.
- **Secure and Trusted environments**: Where justified, AI Factories should establish secure and trusted (research) environments for both industry and scientific research ensuring the confidentiality and integrity of data.

d. Connectivity

AI Factories should ensure a high-bandwidth, low-latency secure networking to support rapid data transfer between nodes and storage systems. In addition, AI Factories should ensure secure connection to the forthcoming EuroHPC Hyper-connectivity network. Indicative references are described below:

- **High-bandwidth, low-latency internal networking**
- **Hyper-connectivity** (e.g. minimum of 100 Gbps, expandable to 1 Tbps).

e. Software and application development

AI Factories should provide a rich software environment including a ready-to-use set of AI-oriented tools (e.g., Pytorch, TensorFlow, etc.) with clear use-cases and examples for efficient use at large-scale, enabling new users to adapt quickly to the environment, as well as to facilitate the use of containerized workloads and workflows. It should be noted that most software tools at the core of AI development and execution are open source and should be supported; otherwise, AI Factories should establish adequate licensing mechanisms.

[12] Common European Data Spaces | Shaping Europe's digital future (europa.eu)

f. User Support for national users and users from the EuroHPC Participating States

Each hosting entity should present their foreseen HPC/AI professional support plan, describing the range of support activities to be offered and provided to users. This may include providing guidance for using the HPC environment, adapting the computational tasks associated to the training and fine-tuning of the models and related inference activities to the HPC environment. User support should be primarily targeting MLOps (machine learning operations). For example, users support activities should include assessing the HPC needs of the users' tasks, providing guidance on missing elements for implementation in HPC environments, parallelization techniques for optimising the memory and computing usage of the hosting supercomputer to speed up (pre-) training, fine-tuning the models for specific datasets and tasks (training or inference), or optimising the final model for efficient deployment and use. The number of required FTEs should be well justified, and the user support team should provide a well-functioning service (below 4h response time where possible).

Applicants should also describe the way they plan to serve public and private users from the EuroHPC Participating States. Such users shall be granted the share of EU's access time to the AI optimised supercomputers and AI Factory services. For such users, hosting entities should propose an appropriate access policy that respects a number of conditions for access (such as for example those in relation to the handling of sensitive information, security, confidentiality, unethical use, etc.).

AI services, including User Support, should be provided in a consistent and professional manner following industrial standards.

g. Co-working and entrepreneurial AI Factory Hubs

Applicants would need to provide a plan for making available physical facilities located nearby or associated to the foreseen AI Factory, such as sufficient large and well-adapted co-working spaces, possibly complemented by virtual working spaces. These will serve startups and SMEs, scientific communities/ talented students and HPC/AI support teams, as well as incubators and accelerators to meet and work on common ideas and projects and get access to capital and to community support that are critical to developing the AI ecosystem.

Hosting entities should also include and/or identify the availability of a physical campus facility located nearby or associated to an AI Factory for hosting talented AI students working or trained in the AI Factory. Such facility would stimulate the relationship between the AI Factory and the local universities to create an environment that can attract the necessary talented human capital and build vibrant, attractive, and dynamic communities of practice along the AI Factory region.

h. Skills/Education

Hosting entities should present a comprehensive AI Factory Skills Plan outlining the skills required for the AI stakeholders they intend to target/serve and how to achieve them. This plan should include the offer of a structured training program highlighting relevant courses, activities, and learning pathways tailored to meet the diverse needs of potential users. Similarly hosting entities should convincingly demonstrate that they have direct access to the necessary human capital and talent and, provide a strategy as to how they intend to collaborate and engage with universities, research centers and other training providers to train and equip students at all levels with the necessary in-demand AI skills. The availability of adequate training facilities (such as for example small GPU sandboxes) at universities or research centers could help them attract and train talent.

Hosting entities should demonstrate capacity to put in place training on advanced subjects such as AI for HPC, Deep Learning, AI Programming environments, etc. Additionally, they should show extensive experience in using different delivery modes to provide advanced training in subject areas that require intensive hands-on experience (on-site, online, hybrid) and capacity to deliver a variety of training actions other than courses such as workshops, hackathons, summer-schools, etc. It is crucial that Hosting Entities also demonstrate the capacity to collaborate with other institutions to deliver training. It will be the responsibility of each AI Factory to design and present a robust and comprehensive set of training/education actions to be implemented.

i. Engagement/ Interacting with the AI community

AI Factories should professionally engage with and serve the broader AI community – from academia and research institutions, to startups, SMEs, and industry – liaising with existing initiatives like TEFs, EDICs, EDIHs and National HPC Competence Centres. AI Factories need to identify the main stakeholders at regional and national level and establish connections through networking events and conferences, sharing knowledge and working together on joint projects. Strategic formal partnerships, talent exchange, and joint initiatives can further strengthen these collaborations. It should be noted that national and local ecosystems should be the starting point for building AI Factories. The organisation and coordination of AI, data and HPC initiatives at the European level is important and ensuring to avoid national silos.

Hosting entities may consider the use and support of external professional service companies to optimise their offering and engagement with the AI ecosystem.

j. AI Factories networking

AI Factories should establish a collaborative framework to ensure effective networking and resource optimisation among themselves (e.g., knowledge sharing, specialisation, assets reutilisation, support, training, staff exchange, etc.). The collaboration between AI Factories is very important to enable a thriving European AI ecosystem. This activity will be developed more extensively at a later stage when several AI Factories are operational.

A particular collaboration use case are HPC/AI projects spanning across two or more AI Factories, where users should have a homogeneous end-to-end experience. The collaboration framework must envisage different formal and informal collaboration mechanisms, including the allocation of resources for this purpose, in order to benefit from synergies and avoid duplication of efforts across the ensemble of AI Factories.

k. Developing trustworthy AI

The AI Factories will cooperate with the AI Office and the TEFs to develop robust guidelines and standards for AI development within AI Factories, aligned with the principles and requirements of the AI Act. These guidelines should cover among other, areas such as data protection, transparency, and accountability. This will help create a unified approach to AI development across Europe and different entities and promote trustworthiness and compliance.

The AI Factories will furthermore work closely with the Testing and Experimentation facilities (TEFs), and the national AI supervision agencies, to test and validate AI solutions developed in the AI Factories to ensure they are considered trustworthy and compliant with the AI Act and robust enough to be used in real world settings.

ANNEX II

AI OPTIMISED SUPERCOMPUTERS FOR AI FACTORIES

It becomes clear that AI Factories need **to deploy timely so that an AI dedicated supercomputing and service infrastructures for Europe's AI start-up and research ecosystem can be operational**.

Three complementary tracks can be considered:

"AI Factories" Track

This track is foreseen for those Hosting Entities that are already hosting a EuroHPC Supercomputer which can demonstrate enough computing resources for training large scale, general-purpose artificial intelligence models and emerging artificial intelligence applications can be appointed as AI Factory.

This track will be implemented through a permanently Open EuroHPC JU Call for Expression of Interest of Hosting Entities to appoint existing EuroHPC Supercomputing systems as an AI Factory. The hosting entity commits to undertake AI Factories activities (i.e., the full range of AI factory services).

Further to the appointment of an existing EuroHPC Supercomputing system as an AI Factory, an implementation grant may be awarded to cover for the AI Factories activities (i.e., services). An amendment to the existing Hosting Agreement should be introduced.

Upgraded AI Optimised Supercomputer Track

This track is foreseen for those Hosting Entities that are willing to upgrade their current EuroHPC supercomputer towards an AI Factory.

This track will be implemented through permanently Open EuroHPC JU Call for Expression of Interest of Hosting Entities to deploy and operate an AI Factory (Upgrade supercomputer to AI + AI Factory (Services, Skill development, User support)).

Further to the selection of Hosting entities, a Call for Tender (e.g., procurement) for the acquisition of the upgrade will be launched in addition to one accompanying grant to cover for the AI Factories activities (e.g. services). The existing grant for operational costs will be adapted in consequence. An amendment to the existing Hosting Agreement should be introduced.

New AI Optimised Supercomputer Track

This track is foreseen for those Hosting Entities that are willing to acquire a new AI Factory optimised Supercomputer.

Permanently Open EuroHPC JU Call for Expression of Interest of Hosting Entities to deploy and operate an AI Factory (AI new system + AI Factory (Services, Skill development, User support).

Further to the selection of Hosting entities, a Call for Tender (e.g., procurement) for the acquisition of the new supercomputer will be launched in addition to 2 accompanying grants to cover for the operational costs of the supercomputer and another one to cover for the AI Factories activities (e.g. services).

It should be noted that these 3 AI Factories Implementation tracks can be implemented in parallel.

ANNEX III

AI Ecosystem Key Features

Key Feature	Key Feature Description	How address it
AI optimised supercomputers	• Is the application developed around an AI optimised supercomputer (existing, upgraded, or new)?	Provision by the Applicants of the description of an AI-optimised supercomputer.
National AI Strategy	• To what extent the establishment and deployment of an AI Factory is linked and contributes to the implementation of the national AI strategy of the hosting country/countries of the hosting consortium?	Provision by the Applicants of the description of the National AI Strategy or equivalent, clearly showing the strategic character of the AI Factory proposal. **NB:** In the absence of a formal national AI strategy, applicants will need to describe the strategic national (or Consortium) character of their AI Factory approach.
National Data Policies	• Is there a current National Data Policy enabling the access to large datasets, availability of knowledge corpus, etc., and if not, is there a plan included describing how the proposal will make available large data sets to the AI Factory ecosystem? • Does the proposal include a plan on how to implement policies facilitating the access to open data and proprietary data (including if necessary different fee schemes depending on the use of data for training/fine-tuning/inference)?	Provision by the Applicants of the description of: 1. National Data policy or equivalent. 2. Meaningful implementation policy for access to large data sets **NB:** the access to available "data" is key to facilitate the functioning of any AI Factory.
Access Policy	• Does the proposal include an AI user-friendly national access policy?	Provision by the Applicants of a description of the access policy to the nationally owned computing time of the EuroHPC supercomputer.

		NB: This is an essential requirement for a part of the application on an AI Factory proposal to provide
Stakeholder participation	• Does the application include a plan on how to attract key national AI stakeholders?	Provision by the Applicants of a description of a convincing plan for attracting such key AI stakeholders. **NB:** This is an essential requirement for an AI Factory.
AI Ecosystem needs and challenges	• Does the proposal describe its strategic focus industrial / application sectors and how it would help develop further the AI ecosystem in these sectors? • Does the proposal include any plans for provision of cloud solutions?	Provision by the Applicants of the description of the key industrial/application sectors as well as of the key obstacles to overcome to further develop the AI innovation ecosystem in these sectors. Provision by the Applicants of any internal or external cloud solutions to bridge the needs towards an end-to-end computing continuum. **NB:** The identification of the above is essential for justifying the need of building an AI Factory that corresponds to the strategic national priorities.
Strategy for startups and SMEs	• Does the proposal include plans for linking to an existing or developing a new national/regional strategy for helping investment in the AI startups and SMEs?	Provision by the Applicants of the description of any plans they have on linking to an existing or developing a new investment strategy for AI start-ups and SMEs. **NB:** While not an essential requirement for an AI Factory, it would help a lot to further grow the national AI innovation ecosystem.
KPIs	• Does the proposal include key performance indicators (KPIs) and targets to measure the contributions to the	Inclusion by the Applicants of a set of meaningful KPI indicators and realistic targets.

	success of the AI Factory and associated AI ecosystem?	**NB:** These are critical to monitor progress and identify where/when needed corrective action.