

EuroHPC JU AI and Data-Intensive Applications Access call

Terms of Reference

Introduction

The EuroHPC Joint Undertaking (JU) enables the coordination of efforts and the sharing of resources at the European level with the objective of deploying a world-class High Performance (HPC) infrastructure and a competitive innovation ecosystem in supercomputing technologies, applications and skills in Europe.

The EuroHPC JU is acquiring supercomputers and quantum computers (the EuroHPC supercomputers) which are located at and operated by supercomputing centres (Hosting Entities) in the Union. The Joint Undertaking manages the Union's access time (from 35% up to 50% of their total capacity) of these supercomputers. Access time is allocated to European scientific, industrial and public sector users, matching their demanding application requirements, according to the principles stated in the EuroHPC JU Council Regulation¹.

Access time is allocated to eligible European users according to the principles stated in Article 13 of the EuroHPC JU Council Regulation¹. Researchers from academia, research institutes, public authorities and industry established or located in an EU Member State or a country associated with <u>Horizon 2020</u>, are eligible to apply for access to EuroHPC JU resources.

AI and Data-Intensive Applications Access call will allocate resources from EuroHPC systems offering optimal hardware for Artificial Intelligence applications; for example compute partitions of EuroHPC systems with accelerators (i.e. Graphics Processing Units). Up to 20% of total time from these systems may be allocated to these projects, subject to allocation demands from other access modes for these systems in a given period (e.g. allocation demands for Strategic Initiatives or Urgent applications).

Proposals will be granted access on a first-come-first-served basis until the allocated resources for the specific cut-off have been depleted, provided the evaluation process is passed successfully.

The proposals submission is done via the Peer-Review portal available at <u>https://access.eurohpc-ju.europa.eu</u>.

Further details on the standard application procedure can be found on EuroHPC JU Call for Proposals AI and Data-Intensive Applications Access mode.

¹ <u>Council Regulation (EU) 2021/1173 of 13 July 2021 establishing the European High Performance Computing Joint Undertaking (OJ L 256, 19.07.2021, p. 3–51)</u>

1 Scope of the EuroHPC JU AI and Data-Intensive Applications Access call

The EuroHPC JU AI and Data-Intensive Applications Access aims to support ethical Artificial Intelligence, Machine Learning, and in general, Data Intensive applications, with a particular focus on Foundation Models and Generative AI (e.g. Large Language Models).

The call is intended to serve Industry organisations, Small to Medium Enterprises (SMEs), startups, as well as public sector entities, requiring access to supercomputing resources to perform Artificial Intelligence and Data Intensive activities.

The call focuses on the training of Large Language Models to serve the requirements of European industry and society. Resources are allocated primarily to perform compute-demanding training executions of ethical AI models. Only a limited percentage of an allocation (no more than 10% of the overall allocation) should be dedicated to inference runs of trained AI models.

The call is continuously open with a cut-off every two months that triggers the evaluation of the proposals submitted up to the cut-off date.

The allocations are granted for one (1) year. As a guiding principle, a person acting as Principal Investigator (PI) can, in principle, be awarded access time for only one Access for AI application at any given time.

This access mode specifically targets ethical AI and Data Intensive applications. Traditional computational applicants should apply to the relevant tracks of Extreme Scale and Regular Access modes.

1.1 Requirements

Applications to EuroHPC JU computing resources must demonstrate compatibility of the codes with the EuroHPC JU systems requested or a need for large data training that requires a large amount of CPU/GPU time overall.

The need for EuroHPC JU systems computing performance must be clearly presented in detail in the proposal.

Resources can be requested on a single system only. Amount of resources that can be granted are **limited with pre-defined fixed values** in terms of node hours per application. The resources thresholds can be found on <u>EuroHPC JU Call for Proposals AI and Data-Intensive Applications Access mode</u>.

Principal Investigators (PI) can, in principle, be awarded access time for only one Access for AI application at any given time. However, where a PI has applied for access time for more than one project, awards of additional projects from the same PI can be granted provided that resources are still available in the given call.







Applicants must submit a full application demonstrating:

- the relevance of the application to the call, including the need for Artificial Intelligence and Data Intensive applications;
- the significant impact of the expected results;
- that their application requires the use of large allocations both in terms of compute and data storage resources- to reach the objective of their application;
- that the methods, software, and tools are technically adapted to efficiently use the target supercomputer thereby demonstrating the feasibility of the project. Applications should clearly explain why the work cannot be performed on a smaller HPC system. If relevant, applicants may rely on technical data collected via <u>Benchmark Access</u>.
- a project plan, with an adequate time schedule of the expected resource consumption during the lifetime of the project as well as a GANTT chart.
- ethical AI by completing an ethics self-assessment form.

Furthermore, the following applies:

• Generated data and models remain under the ownership of the user. SMEs and startups may use the outcome of these allocations for commercial exploitation.

Projects not covered under the AI and Data-Intensive Applications Access Call:

- Proposals for code testing and optimisation are outside of the scope of this call. A separate call for **Benchmark and Development Access** is continuously open for such purposes.
- Proposals for traditional computational applications for Research and Innovation, Industries and SMEs should be submitted to the relevant tracks of **Extreme Scale** and **Regular Access** modes.

2 Eligibility criteria

Applicants from industry organisations, Small to Medium Enterprises (SMEs), startups, as well as public sector entities and research organisations are eligible to apply as long as:

- a) the organisation is established or located in an EU Member State or in a country associated to the Digital Europe Programme or to Horizon Europe, or where applicable, organisations residing, established or located in an EU Member State or in a country associated to Horizon 2020;
- b) the Principal Investigator has an employment contract in the organisation at the time of proposal submission and valid for at least 3 months after the end of the allocation period; and
- c) for what concerns access to commercial companies and Small and Medium Enterprises (SMEs), the relevant Horizon 2020 <u>rules of participation</u> shall be applied.

Only proposals with a civilian purpose will be eligible to participate in EuroHPC JU calls for proposals. Only proposals written in English will be eligible. **Double-awarding is not allowed**; proposals already granted resources as part of the EuroHPC JU Access calls, and currently running, will be rejected.



Proposals will be evaluated by experts, based on the award criteria of 'excellence', 'innovation and impact' and 'quality and efficiency of the implementation'. In detail, the following aspects are considered during the evaluation for each evaluation criterion:

a. Excellence

This criterion aims to evaluate the quality and merit of the project through the following dimensions:

- Clarity and relevance of the objectives;
- Soundness of the concept;
- Credibility and appropriateness of the proposed methodology;
- The methodology and the technical maturity of the project;
- The expected outcomes;
- Demonstrates the need for supercomputing resources;

b. Innovation and Impact

This criterion intends to assess the innovative nature, the potential impacts and contributions of the project. It evaluates to what extent the proposed work is beyond the state of the art, and demonstrates innovation potential (e.g., ground-breaking objectives, novel concepts and approaches, new products, services or business and organizational models). The following dimensions are considered:

- Originality and novelty of the proposed objectives and project;
- Innovative nature of the proposed project;
- Contribution of the project to the advancement of the state of the art and potential innovations;
- Impact of the project results on societal, economic, and/or technological dimensions, including expected impact on trustworthy AI applications.

c. Quality and efficiency of the implementation

This criterion is intended to evaluate the quality and feasibility of the project work plan to deliver the project successfully. Proposals should demonstrate a balanced distribution of resource utilization during the lifetime of the project, ensuring a stable consumption of the total awarded time during the awarding period.

The following dimensions are considered:

- Feasibility of the work plan;
- Alignment between the resources requested and the objectives of the project;
- Appropriateness of resource allocation schedule to successfully complete the project;

• Appropriateness, quality and skills of the team members to perform the proposed objectives and tasks and accomplish the proposed schedule.

3.1 Scoring principles

Experts score each award criterion on a scale from 0 to 5 (half-point scores may be given):

- 0 Proposal fails to address the criterion or cannot be assessed due to missing or incomplete information.
- **1** Poor. The criterion is inadequately addressed or there are serious inherent weaknesses.
- 2 Fair. The proposal broadly addresses the criterion, but there are significant weaknesses.
- 3 Good. The proposal addresses the criterion well, but a number of shortcomings are present.
- 4 Very good. The proposal addresses the criterion very well, but a small number of shortcomings are present.
- 5 Excellent. The proposal successfully addresses all relevant aspects of the criterion. Any shortcomings are minor.

The maximum overall score is thus 15. The minimum threshold for each criterion is 3 and the overall threshold for the sum of all criteria is 10. All criteria are equally weighted (1.0).

Proposals will be granted access on a first-come-first-served basis until the allocated resources for the specific cut-off have been depleted, provided the peer-review process is passed successfully.



The Principal Investigator shall lead the project and is expected to be an essential participant in its implementation. The PI will have the overall responsibility for the management of the project and interactions with EuroHPC JU. The applicants should make sure that the contact details for the PI are consistent in the different forms to be completed and that **all e-mail addresses used are professional e-mail addresses**.

The usage of EuroHPC JU resources needs to be acknowledged for all data produced through EuroHPC JU allocations, both in publications and when depositing the data to other infrastructures.

The PI commits to:

- a) Provide to EuroHPC JU a **final report within 3 months** of the completion of an allocation, using the proper EuroHPC JU <u>template</u>, with the results obtained through the access to the EuroHPC JU systems, as well as qualitative feedback on the use of the resources.
- b) **Acknowledge** the role of the HPC Hosting Entity and EuroHPC JU in all publications which include the results above mentioned. Users shall use the following wording in such acknowledgement in all such papers and other publications:

"We acknowledge EuroHPC JU for awarding the project ID EHPC-AI-XXXXAXX-XXX access to *[resource-name hosted by at site]*"

Use as many instances of the pattern [resource-name hosted by at site] as the number of systems awarded via EuroHPC JU.

Respecting the words in bold above is particularly important since EuroHPC JU will use this word pattern when searching for bibliographic references in scientific articles. In case additional resources have been used, the acknowledgement should include a clear breakdown of which part of the work was performed using EuroHPC JU resources.

- c) **Allow** EuroHPC JU to publish the report mentioned in section (a) above after one year from the termination of the allocation period.
- d) Contribute to EuroHPC JU dissemination activities and other EuroHPC JU events. Selected awardees are expected to contribute to and attend such events at least once over the two-year period, starting from the end of the allocation period. Awardees will also be expected to reply favourably, when asked to be interviewed for EuroHPC JU publications and/or send visualizations or other materials for promotional purposes.

Access to EuroHPC JU resources is for **is free of charge** provided that the eligibility criteria and terms of access described herein and in the online Application Form are fulfilled/respected. If this differs from the terms of access that the relevant Hosting Entity may have in place, it is the terms of access of the relevant Hosting Entity that will prevail.







Users will not hold liable EuroHPC JU or the relevant Hosting Entity, including their Directors and staff, with regard to any claim and expense arising out of the use of the resources.

From the start to the end of the access period, the applicant should direct questions and requests for support to the user support of the HPC Hosting Entity(s) where resources have been allocated.

Applicants must inform promptly the EuroHPC JU and the Hosting Entity where the resources are allocated of any changes to a successful proposal, namely a decrease in the amount of resources needed or on the distribution of the usage of the resources within the agreed time plan with the Hosting Entity.

The Applicants may request an extension of their allocation of up to 3 months in the event that they were not able to consume the assigned resources and complete their project in the allotted time. In such cases, the PI should submit a formal request to EuroHPC JU PRO at least 1 month before the conclusion of the project. If the underutilised time is more than 20% then the excessive time is removed from the total remaining allocation. No allocation of additional resources is possible.

5 Process details

5.1 How to apply

All proposals must be submitted via the Peer-Review portal available at <u>https://access.eurohpc-ju.europa.eu</u>.

All proposals must be fully completed and submitted by the closing cut-off date and time in order to be considered for evaluation for that submission period. The submission platform will not accept applications for a dedicated cut-off date that are submitted after the respective cut-off date and time have passed. All proposals submitted after the cut-off date and time will be considered for assessment in the next cut-off period. In the case of technical difficulties, the decision of EuroHPC JU as to whether an application can be accepted is final.

Applicants are advised to make sure, that they submit proposals as early as possible before the given deadline, in order to ensure that all mandatory fields are completed and submission is accepted.

Further details on the standard application procedure can be found online in the following links: "How to Apply" and "Information for Applicants" in the "AI and Data-Intensive Applications Access - Full Call Details" document.







5.1.1 Applications

All proposals consist of 2 parts: an online form and the **'Project scope and plan**' document, to be submitted via the **Peer-Review** portal available at <u>https://access.eurohpc-ju.europa.eu</u>.

The template of the '**Project scope and plan**' document (Uploaded on the online application form as a .pdf) must be <u>carefully respected</u> (Headings, length, tables and figures). **Proposals that do not follow the template or that are incomplete will be administratively rejected and will not be further evaluated.** The Peer-Review team is available to answer questions by email (<u>access@eurohpc-ju.europa.eu</u>).

All mandatory fields of the online application form must be completed before it can be submitted. Please note that <u>only submitted proposals will be put forward for peer-review</u>.

The template documents for these reports are available on the **EuroHPC JU Call for Proposals AI and Data-Intensive Applications Access mode** ("<u>Documents</u>") and must be **carefully respected**.

5.2 Peer-Review assessment procedure

The assessment procedure (Peer-Review process) abides to the EuroHPC JU Peer-Review principles stated in **"The Peer-Review Process**". The Peer-Review process encompasses 3 phases:

- Administrative Check. Proposals not complying with EuroHPC JU requirements and eligibility criteria will be rejected at this stage and will not continue to the next phase. During this phase, applicants may be contacted by the Peer-Review Office in case of concerns regarding potential clerical errors.
- 2. **Technical Assessment.** Proposals will be technically reviewed by technical experts of EuroHPC JU Hosting Entities, who will assess the suitability of the application to run in the indicated system. Proposals rejected by the technical experts, will not be awarded.
- 3. **Experts Evaluation.** Proposals will be peer-reviewed by recognized independent experts against the criteria in Section 3. Proposals ranked under the quality cut-off threshold will not be awarded, even if there are resources available on the systems.

Upon successfully passing all the above steps, allocations are made on a first-come-first-served basis until the resources reserved for the specific cut-off are exhausted.

In case a proposal is positively evaluated by technical and peer-review experts and is recommended to be awarded but it cannot be allocated due to lack of resources on the desired partition (first-come-first-serve principle), EuroHPC JU reserves the right to allocate the proposal in the next cut-off period with allocation







priority. This needs to be indicated by the Applicant in the online form. The same proposal can be moved from one cut-off to another just once.

Successful applications are notified by email and are contacted by the technical teams of the assigned Hosting Entity for the onboarding procedure. Prospective users will be requested to sign an Acceptable User Policy before accessing the system.

5.2.1 Right to appeal

Applicants whose proposal has not been awarded, may appeal to the decision by sending a formal letter by email to the EuroHPC JU (access@eurohpc-ju.europa.eu) within 15 days from the date of reception of the rejection decision. The appeal letter should clearly state the reasons why the applicant considers that the evaluation or eligibility check results were incorrect, by referring to the rejection arguments listed in the response letter. The letter should not be a mere resubmission of the initial proposal.

A review committee will assess the reasons of appeal and will respond to the applicant within a maximum period of 1 month from the date of reception of the appeal. In case of eligibility rejection, the committee will recheck the reasons of rejection while in case of low ranking, the committee will assess whether there are grounds for proposal re-evaluation. In the latter case, a new evaluation committee will be assigned to perform the review.

Should the committee accept the grounds of redress, in case of eligibility rejection, the application will be included for evaluation in the next cut-off period. In case of low ranking, the application will be prioritized for allocation in the resources available for the next cut-off period.

6 Tips and examples

This section includes a few tips and examples of common mistakes or misunderstandings in the preparation and submission of proposals:

- a) **Submission deadline.** A team faces last-minute problems, not related to the submission system, in the submission of their proposal and is not able to submit it completely before the deadline. The application is not considered for the current cut-off.
- b) **Submission completeness.** An application is received incomplete, i.e. missing documents or documents with missing sections. The application is administratively rejected, and it will not be evaluated.
- c) **Application exceeding limits.** A research team submits a proposal exceeding the page limits. The exceeding pages will not be considered as part of the application. Reviewers will be instructed not to consider the exceeding pages, and this may imply that the application is administratively rejected.

- EuroHP Joint Undertaki
 - d) **Technical readiness**. A research team submits an application lacking information regarding the technical feasibility of the project, assuming that they will be able to provide this data during the evaluation of their proposal. The application is administratively rejected since this data is mandatory at the time of submission.

7 Contacts

For any queries related to applications, please contact: <u>access@eurohpc-ju.europa.eu</u>