Call for Pilot Foundations:
Helmholtz Foundation Model Initiative

Foundation models are defined as "(…) any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks;" (arXiv:2108.07258v3). Recent examples demonstrated the potential of foundation models, e.g., RETFound, Segment Anything and IBM NASA Geospatial.

The Helmholtz Foundation Model Initiative (HFMI) took the lead in creating AI models that harness large and complex datasets to help address a variety of tasks in several research domains. In May 2024, The Helmholtz Association launched the ambitious initiative to leverage the recent advancements in AI and the rise of "Foundation Models", initially selecting four projects, accompanied by a Synergy Unit that addresses overarching issues.

With this call, we extend our support for highest-impact pilot foundation model projects within the Helmholtz Association. We invite all Helmholtz researchers to participate in this competitive call with their specific existing datasets, domain-focused research questions, and unique foundation model approach. Successful projects will demonstrate a clear path to transform research data into high-impact tools for groundbreaking scientific discovery.

Consortial structure

A consortium of Principal Investigators (PIs) will guide each pilot foundation project, working together to create a high-impact foundation model in the specific research domain. Proposals should demonstrate the expertise of the participating PIs in three aspects: 1) Research domain expertise, 2) AI expertise, 3) HPC expertise. Project teams will comprise data specialists with demonstrated domain expertise, as well as AI specialists with demonstrated core machine learning expertise, and HPC specialists with demonstrated experience translating large computational loads to successful HPC runs. We recommend to assign specific "Domain Lead", "AI Lead", and "HPC Lead" roles among participating PIs.

HFMI follows the idea of the step-wise approach, with Foundation Stage, Readiness Stage, and Visionary Stage, with progression from single-modal data or a few closely related modal data in the foundation stage to multi-modal in the visionary stage. Therefore, every application should explicitly lay out the envisioned evolution between the stages and incorporated data.

More information about the desired consortial structure can be found in the document "Helmholtz Foundation Model Initative" included with this call.

Eligibility

All Helmholtz researchers are eligible to apply. External partners can participate in consortia but are not eligible to receive funding though this scheme.

Duration

Projects should run for up to 36 months, starting early 2025.

Funding scheme

We intend to support two projects with 1,500,000 EUR each towards personnel, equipment, and associated costs. Funds must be spent entirely in 2025.

The contributing centers must underwrite funding for the remaining period starting January 2026, since central funding will not be available after 2025. This commitment must be confirmed by a letter of support signed off by the research center leadership(s) of participating PIs and included in the proposal PDF.

The recommended team structure comprises 8 FTE with domain and/or data experts hosted at centers with relevant domain expertise, AI experts hosted at a center with relevant AI expertise, and HPC experts at centers with relevant HPC expertise and resources.

Selection Criteria

Projects should leverage existing large & complex data sets to build Foundation models to solve several highly impactful, well-defined downstream tasks, with quantifiable success.

Projects will be selected under the premise that building a Foundation Model requires the convergence of leading expertise in 1) the research domain, 2) AI, 3) HPC. In addition, ready access to datasets of the complexity required to give rise to a foundation model.

A consortium of at least two Principal Investigators (PIs) will guide each pilot foundation project: one expert from the respective research area ("Domain Lead"), and one from the AI field ("AI Lead"). In addition, we recommend to include a PI who has HPC expertise ("HPC Lead").

Successful proposals will provide information on:

- Provenance, availability of, and access to complex, high-volume data, and how its extent and complexity are expected to give rise to a Foundation Model.
- Evidence of leading expertise in
    1. the relevant research domain,
    2. relevant AI approaches pertinent to the chosen Foundation Model approach,
    3. leveraging large-scale HPC resources, including the handling of large complex datasets.
- How consortia merge domain, AI, and compute expertise to build domain-leading Foundation Models.
- The amount of compute resources required and a detailed and reliable plan to access them throughout the duration of the project.

In addition, proposals should demonstrate:

- Scientific relevance and the extent to which the model provides solutions to the central questions of our time ([Helmholtz Mission](#)).
- A viable research program including defined objectives and desired results for each project phase, with work plan, milestones, deliverables, including an assessment of risks that could impact the successful completion of tasks, and their mitigation.
- A specification of foundation model approach (e.g. architecture, training & tuning paradigms, etc.), and proof of ability to scale to large-scale HPC.
- A concept for cooperation within the consortium ("[team science](#)"), i.e., how an effective team spirit will be formed in the consortium in the light of the interdisciplinary nature and potential distributed location of consortium partners.
- Qualification and track record of personnel in the research domain, AI/ML, and HPC.
- Assessment of readiness for a timely start of the project.

The proposal template contains additional guidelines for the information that should be provided.

Application Process

The rapid progress in AI commands a timely initiation of projects. To this end, the application process will be implemented as follows:

23.07.2024	Call published
31.07.2024	Q&A event about the call – 10:30-11:30 am CEST
16.09.2024	Submission deadline - 09:00 am CEST
14.10.2024	Selection meeting and decision
Early 2025	Start of projects

The provided template must be used to prepare the proposal.

Submission

Proposals must be submitted as a **single PDF less than 10 MB in size** via email to inkubator@helmholtz.de. Emails larger than 10MB will be rejected by the email system. We will acknowledge the receipt of submissions during normal working hours.

Formal requirements

Submission checklist

Your submission must meet all of the following criteria.

a. Strictly adhere to the proposal template regarding the structure, word and page limits of your proposal, including mandatory appendixes and CVs. In the spirit of fair competition, please refrain from tweaking font and pages sizes.
b. Submitted proposals must be complete at the deadline, i.e. all parts of the application addressed and all requested information provided at the time of submission.
c. Financial planning must be transparent and sound.
d. The proposal PDF must contain a letter confirming the financial commitment by the host center(s) for the project period beyond 2025.
e. If applicable, provide signed declarations or guarantees of matched funds from each Helmholtz center, or letters of support from external partner.

Reporting and other obligations

Projects are expected to integrate in the existing HFMI structure that is coordinated by the Synergy Unit. They will be required to participate in joint dissemination events and monitoring exercised.

HFMI PIs shall prepare reports that demonstrate progress in the project plan (milestones) and results (deliverables) to Helmholtz head office and the Helmholtz Information and Data Incubator. Reports are due annually and at the end of the funding period.

Project leaders commit to contribute to Helmholtz Foundation Model Initiative activities. Projects teams are expected to actively participate in activities coordinated by the Synergy Unit, e.g. in progress or method exchange workshops, the Helmholtz AI and Imaging conference, and other research and dissemination events.

It is mandatory for funded PIs and researchers to acknowledge financial support through the Helmholtz Foundation Model Initiative funding in any related published output.

Funded PIs commit to regularly communicate updates about ongoing related activities, for publication e.g. on the Helmholtz website and/or associated social media channels. The head office may approach PIs for such updates.