The Helmholtz Foundation Model Initiative (HFMI) - Harnessing Foundation Models for Science and Society

Eirini Kouskoumvekaki and **Dagmar Kainmueller** *Max Delbrück Center, Berlin, Germany* {eirini.kouskoumvekaki,dagmar.kainmueller}@mdc-berlin.de
Stefan Kesselheim *Jülich Supercomputing Center, Forschungszentrum Jülich, Germany* s.kesselheim@fz-juellich.de
Stefan Bauer *Helmholtz AI, Munich, Germany* stefan.bauer@helmholtz-munich.de
Fabian Isensee *German Cancer Research Center, Heidelberg, Germany* f.isensee@dkfz-heidelberg.de

Abstract

The Helmholtz Foundation Model Initiative (HFMI) develops open, interdisciplinary foundation models (FMs) to accelerate science and deliver societal benefit. Building on Helmholtz's unique strengths, namely extensive datasets, world-class infrastructure including Europe's first exascale supercomputer JUPITER, and interdisciplinary expertise, HFMI demonstrates how FMs can address grand challenges across climate, health, energy, and life sciences. Seven pilot projects launched in 2024–2025 illustrate this breadth, supported by a Synergy Unit that ensures scalability, interpretability, and knowledge exchange. Guided by FAIR and open science principles, HFMI democratizes AI access and fosters collaboration.

Keywords: Foundation Models, AI, Open Science, Supercomputing, Interdisciplinary Research

1 Introduction

Artificial intelligence is transforming the way science is conducted. Among the most powerful advances are Foundation Models (FMs); AI systems trained on massive and diverse datasets with unprecedented computing power. Unlike traditional machine learning models designed for narrow, domain-specific tasks, FMs acquire a broad knowledge base during pretraining. This allows them to be flexibly adapted for a wide variety of downstream applications with minimal fine-tuning, making them uniquely suited to address the complex, data-intensive challenges facing modern science.

The potential impact of FMs extends across all major research domains: in climate science, they can integrate atmospheric, oceanic, and cryospheric data for improved forecasts; in health, they can accelerate diagnostics, drug discovery, and personalized treatments; in energy, they can enable faster discovery of

sustainable materials; and in the life sciences, they can deepen our understanding of fundamental biological processes. By lowering the technical barriers to advanced AI use, FMs have the capacity to democratize access to cutting-edge research tools, empowering scientists regardless of discipline or resources.

With more than 45,000 staff members, including 23,000 active researchers, the Helmholtz Association provides an ecosystem of scale and diversity that few institutions worldwide can match, and offers unparalleled conditions for foundation model (FM) development by combining three essential strengths: it maintains extensive repositories of scientific data across domains such as health, climate, and energy; it provides world-class infrastructure, including supercomputers and advanced data platforms; and it unites interdisciplinary expertise, bringing together AI specialists and domain scientists.

The Helmholtz Foundation Model Initiative (HFMI) (https://hfmi.helmholtz.de/) coordinates these assets under one umbrella, fostering collaboration, efficient resource use, and maximum impact across all Helmholtz centers. HFMI is the result of a broad community driven discussion process. The initiative is bringing together leading minds in AI research and domain science with the goal to provide instruments answering the grand challenges of our time. Key partners in HFMI include Helmholtz Centers across six research fields, each contributing domain-specific data and expertise. Additionally, the well-established Helmholtz Framework Information and Data Science with its five platforms – the Helmholtz Metadata Collaboration (HMC), Helmholtz Imaging, the Helmholtz Federated IT Services (HIFIS), the Helmholtz Information and Data Academy (HIDA) and the Helmholtz Artificial Intelligence Cooperation Unit (Helmholtz AI) – play a crucial role in supporting data curation, providing computational infrastructure, enabling AI integration and supporting training and translation of first results into the broader Helmholtz community. These partnerships ensure coordinated implementation, with shared resources like supercomputers and datasets driving the development of foundation models. Collaborations with academic institutions, industry partners, and international research networks extend HFMI's reach, fostering interdisciplinary innovation and amplifying its global impact.

HFMI has been mandated a pioneering mission: to develop open-source, interdisciplinary foundation models that not only accelerate scientific progress but also promote fairness, inclusivity, and transparency in AI. Through its pilot projects and overarching Synergy Unit, HFMI demonstrates how FMs can be harnessed to tackle grand societal challenges, while ensuring that the results – data, models, and code – remain openly available for the global research community.

2 Supercomputing Power: JUPITER and HAICORE

Foundation Models are only possible with immense computing capacity. Helmholtz provides this through JUPITER, the first European exascale computer at Jülich Supercomputing Center, capable of exceeding one quintillion calculations per second. JUPITER is poised to be one of the most powerful AI supercomputers

worldwide, enabling the training of large-scale FMs with unprecedented speed and efficiency. This flagship system is complemented by Helmholtz AI computing resources (HAICORE), a distributed infrastructure of powerful clusters that expands access to advanced computing across Helmholtz centers. Together, JUPITER and HAICORE provide not only a first-class research environment but also a strategic framework: ensuring that resources are used efficiently, sustainably, and with guaranteed scientific impact.

3 HFMI Pilot Projects and Synergy Unit

During two Helmholtz-wide calls in 2024, seven pilot projects and one cross-cutting Synergy Unit were selected for funding, involving scientists from fourteen Helmholtz Centers. A key unique feature of the HFMI is the collaboration between AI experts and domain scientists within these projects, ensuring that not only do the models make a meaningful contribution to research, but also that the results withstand empirical scrutiny. Over an initial funding period of three years, the projects will receive funding of 24 million Euros, with an additional 11 million Euros invested in expanding the necessary infrastructure.

Table 1: Overview of HFMI pilot projects, the cross-cutting Synergy Unit and their main research goals.

Project	**Domain**	**Main Goal**
3D-ABC[1]	Environment	Quantify and visualize the global carbon budget of vegetation and soils using multi-source observational data.
THRP[2]	Health	Create a foundation model from large-scale 3D radiology data to enhance diagnostics and personalized medicine.
HClimRep[3]	Climate	Build a foundation model integrating atmosphere, ocean, and sea ice data for high-resolution climate predictions.
SOL-AI[4]	Materials	Accelerate the discovery and optimization of photovoltaic materials by integrating diverse datasets.
VirtualCell[5]	Life Sciences	Develop a digital twin of a cell using multi-omics and spatial data to advance biomedical research.
AqQua[6]	Marine Biology	Build a foundational pelagic imaging model from billions of marine images to monitor biodiversity and carbon cycling.
PROFOUND[7]	Biomedicine	Model protein dynamics beyond static structures to enable smart therapeutics and molecular machines.
Synergy Unit[8]	Cross-cutting	Address scalability, interpretability, and knowledge exchange across all projects.

Each of the funded projects demonstrates the transformative potential of FMs across disciplines, as illustrated in the table above. Every project is built around a core group and supported by a broader network of institutions that contribute data or specialized expertise. These networks are uniting the global plankton community, as in AqQua; fostering collaboration with industry, as in THRP; or advancing existing scientific models, as in HClimRep and 3D-ABC, through partnerships with ESA and NASA.

The Synergy Unit serves as the connective element of HFMI, addressing interdisciplinary questions and fostering knowledge exchange across individual pilot projects. Its central mission is to accelerate research on Foundation Models across disciplinary boundaries while ensuring that the initiative's long-term impact contributes to the common good. The Synergy Unit focuses particularly on methodological challenges that emerge across the foundation model pipeline, such as scaling behavior in large-scale optimization Filatov [2025b,a] or model generalization under diverse kinds of distribution shifts Winklmayr [2025].

4 Impact

All partner Helmholtz centers play leading roles in their respective research fields, ensuring a strong alignment between HFMI projects and the needs of the scientific communities they serve. For example, the initiative has fostered discussions on interpretable AI and the evaluation of AI model outputs. With the call UNLOCK for funding of benchmarking projects that was announced in spring (https://www.helmholtz.de/en/research/current-calls-for-applications/article/helmholtz-benchmark-projects-call-unlock-2025/), Helmholtz has taken up these discussions and is promoting the development of improved benchmarks to enable more reliable and effective AI models.

Furthermore, each Helmholtz Center is well-integrated within its region, promoting robust local partnerships and collaboration opportunities. This connectivity ensures that the advantages of our AI resources extend beyond the scientific community, reaching societal stakeholders and addressing real-world challenges. By promoting open-source tools, FAIR data, and knowledge exchange, we democratize access and empower researchers globally, fostering fairness, inclusivity, and broad participation in AI innovation. This large-scale, open-access effort enhances research efficiency, reduces costs, and drives innovation with tangible benefits for society, such as improving healthcare, advancing clean energy, and supporting climate resilience.

[1] https://www.3d-abc.ai/
[2] https://human-radiome-project.de/
[3] https://hclimrep-project.de//
[4] https://hfmi.helmholtz.de/pilot-projects/sol-ai/
[5] https://hfmi.helmholtz.de/pilot-projects/virtualcell/
[6] https://www.aqqua.life//
[7] https://hfmi.helmholtz.de/pilot-projects/profound/
[8] https://hfmi.helmholtz.de/synergy-unit/

5 Conclusions

HFMI positions Helmholtz as a leader in open, interdisciplinary AI. By uniting vast data resources, powerful computing, and scientific expertise, we aim to push the boundaries of research, democratize AI, and ensure its benefits are shared widely for science, for innovation, and for society.

Acknowledgements

The authors would like to acknowledge the support of the Helmholtz Association within the framework of the Helmholtz Foundation Model Initiative Synergy Unit project.

References

O. Filatov. Optimal scaling needs optimal norm. In *arXiv:2510.03871*, 2025a.

O. Filatov. Time transfer: On optimal learning rate and batch size in the infinite data limit. In *arXiv:2410.05838*, 2025b.

C. Winklmayr. Phenobench: A comprehensive benchmark for cell phenotyping. In *arXiv:2507.03532, MICCAI 2025*, 2025.