

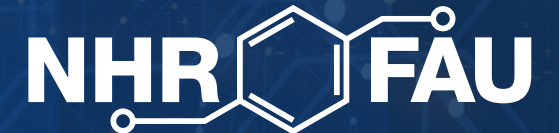
BayernKI - Ressourcen und Support für die bayerische KI- Forschung

Prof. Dr. Gerhard Wellein, Dr. Nicolay Hammer, Dr. Thomas Zeiser
Martin Mayr, Benjamin Geißler, Dr. Mares Barekzai

Bayerische Zusammenarbeit LRZ + NHR@FAU Hand in Hand für Hochleistungsrechnen & KI in Bayern



Leibniz-Rechenzentrum
der Bayerischen Akademie der Wissenschaften



Friedrich-Alexander-Universität
Erlangen-Nürnberg



Agenda



- BayernKI – Konzept
- BayernKI – Stand des Projekts
 - Exkursion Helma
- Niederschwelliger Zugang
- BayernKI – Ausblick

ki.bayern.de/wissenschaft oder ki-in-bayern.de

KI-Offensive Bayern

Bayern gestaltet die revolutionären Veränderungen, die Künstliche Intelligenz (KI) in sämtlichen Bereichen der Gesellschaft mit sich bringen wird, aktiv mit. Mit über 130 neuen KI-Professuren und weiteren Maßnahmen aus der Hightech Agenda Bayern hat der Freistaat beste Voraussetzungen für ein erfolgreiches bayerisches KI-Ökosystem geschaffen. Ein kraftvolles Paket von eng miteinander verzahnten Projekten verleiht ihm jetzt zusätzlichen Schub:

1. Bayern baut eigene KI-Rechnerinfrastruktur für die Wissenschaft: Bayerns KI-Expertinnen und Experten an den Hochschulen brauchen für ihre Forschungsprojekte Zugang zu ausreichender Rechenkapazität. Daher baut der Freistaat eine Bayerische KI-Rechnerinfrastruktur für die Forschenden an den bayerischen Hochschulen auf.

Ab 2024 werden am Leibniz-Rechenzentrum in Garching (LRZ) und dem Regionalen Rechenzentrum Erlangen (RRZE) starke KI-Cluster mit Prozessoren neuester Bauart errichtet. Hierfür stellt der Freistaat im Doppelhaushalt 2024/25 im Rahmen der Hightech Agenda bis zu 55 Mio. Euro bereit.

Begleitend werden die beiden Rechenzentren ein niederschwelliges Zugangsverfahren für dieses neue Angebot entwickeln. Auch die KI-Benutzerbetreuung sowie die methodische Beratung und Unterstützung von KI-Anwendern an den Hochschulen werden das LRZ und das RRZE weiter ausbauen.

<https://www.bayern.de/bericht-aus-der-kabinettsitzung-vom-6-februar-2024/>

BayernKI – Konzept

HPC-/KI-Versorgung in Bayern

- **BayernKI: Bayernweite KI-Grundversorgung**

- Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften
- Zentrum für Nationales Hochleistungsrechnen Erlangen



- „Basismodell“ → Projektzugang „only“

- Projektversorgung – **Antragsbasiert** (deutschlandweit)

- „Tier-2“: NHR@FAU – Projekte mit hohem Bedarf → NHR
- „Tier-0/1“: LRZ – Projekte höchster Leistungsklasse → GCS

- Einfacher, **niederschwelliger Zugang zu den BayernKI Ressourcen**
 - Schlankes & schnelles Zugangsverfahren
 - Projektlokale Verwaltung der Zugänge und Projektressourcen
 - **Nachgängige Berichtspflicht/Begutachtung** von Projekten
 - Stoßzeiten-Regelung vor großen KI-Konferenzen
- Nutzungskonzept
 - **Nutzungsberechtigung: Forschende an bay. Univ. & HAWs – kostenfrei!!**
 - Einfache techn. Zugangsverfahren/Authentifizierung (exist. Portale)
 - Basiskontingent von 5.000 GPUh mit möglicher Erweiterung nach Projektbericht

- **Betrieb, User-Support & Training** durch LRZ und NHR@FAU
- Koordination durch **BayernKI Steuergruppe**
 - MD Greiner (StMWK), MR Antretter (StMWK)
 - Prof. Eymann (U Bayreuth), Prof. Groß (TH Nürnberg)
 - Prof. Martin (UTN), Prof. Seidl (LMU, MCML)
 - Prof. Kranzlmüller (LRZ), Prof. Wellein (NHR@FAU)
- Ressourcenvergabe / Nutzungsrichtlinien: Nutzungsausschuss
 - Prof. H.-J. Bungartz (TU München)
 - Prof. A. Hotho (Univ. Würzburg)
 - Prof. D. Merhof (Univ. Regensburg)
 - Prof. T. Bocklet (TH Nürnberg)
 - Prof. M. Botsch (TH Ingolstadt)
 - Prof. J. Grubert (HS Coburg)

Ernannt durch: Universität Bayern e.V. und Hochschule Bayern e.V

LRZ und NHR@FAU als Träger der BayernKI

- **Abgestimmtes Betriebs- und Supportkonzept**
- **Gemeinsames Hardware- und Beschaffungskonzept**
 - Dicht gepackte GPU-Systeme mit Warmwasserkühlung
 - Abgestimmte Investitionsplanung 2025 – 2027
 - Gemeinsame Rahmenvereinbarung in Vorbereitung inkl. Öffnungsklausel für bay. Universitäten und Hochschulen

BayernKI – Stand des Projekts

Aktueller Stand:

- Erste BayernKI-Hardware in Betrieb (120 + 200 NVIDIA H100 GPUs)
- Rahmenvertrag ausgeschrieben und bezuschlagt – Installation Q3/26
- Gemeinsames Support Team aktiv
- Aktive Nutzung durch alle bayerische Universitäten und Hochschulen
- Nutzungsausschuss besetzt
- BayernKI Webauftritt & Zugangsformular: <https://ki.bayern.de/wissenschaft>

	LRZ	NHR@FAU
Installation Q4/2024	120 NVIDIA H100	200 NVIDIA H100 (zusätzlich Kapazitäten auf Alex)
2025 – gemeinsame Ausschreibung für bedarfsorientierte Erweiterung/Rahmenvereinbarung		
Zuschlag erfolgt		
2026	> 180 NVIDIA B200	> 72 NVIDIA B200
2027	??	??

Technologieoffen
Bedarfsorientiert
Warmwasserkühlung
Speichersysteme

Rahmenvereinbarung - BayernKI



- NHR@FAU & LRZ: Europaweite Ausschreibung - Rahmenvereinbarungen für BayernKI
- Zuschlag im November 2025 → Feste Preise bis 11/2026
- Danach Möglichkeit der Verlängerung um 12 Monate (bis zu 3x)
- Bezugsberechtigt: LRZ und alle bay. Hochschulen/Universitäten

- BayernKI / GPU Optionen:

- Lenovo ThinkSystem SC777 V4 Neptune
NVIDIA GB200 NVL4 : 2 x Grace CPU (72 c) & 4 x B200 GPU



Preis-
garantie



- MEGWARE GPU Server mit 8x NVIDIA RTX 6000 BSE
2 x AMD CPU (128 c) & 8 x RTX PRO 6000



HS Hof / HS München
/ Uni Bayreuth:
5x GB200NVL4
2x RTX6000

- Professur für Medizin und Mediztechnik
- Professur für Entrepreneurship
- Projekt nxtAIM
- Institute for Applications of Machine Learning and Intelligent Systems (IAMLIS)

Prof. Dr. Christian Hanshans

Prof. Dr. Sebastian Dünnebeil

Prof. Dr. Fabian Flohr

Dr.-Ing. Martin Häusl

Exkursion: Helma



BAYERNKI



Helma Cluster (NHR@FAU)

MEGWARE/Lenovo AMD / NVIDIA

768 NVIDIA H100/H200 GPUs

- 96 nodes with 4 H100 94 GB HBM2e
- 96 nodes with 4 H200 141 GB HBM3e
- node-local NVME storage
- NDR200 Infiniband

Multi-PB NVMe storage system

BayernKI share: 200 NVIDIA H100 GPUs

Helma	TOP500
#51	06/25
#58	11/25
#73	06/26

Helma – GPUs at work

768 NVIDIA GPUs
(H100 / H200)

#51 weltweit
#3 national
#1 Unis

5 PB NVME Storage
#3 weltweit

120.000 AMD Turin cores
(in der Testphase)



Helma – ein gemeinsamer Kraftakt



Friedrich-Alexander-Universität
Erlangen-Nürnberg



KATHOLISCHE UNIVERSITÄT
EICHSTÄTT-INGOLSTADT



Helma – GPUs dicht gepackt & energieeffizient

Max. Leistung:

~700 kW

~90 kW / Schrank

„Warmwasserkühlung“



Helma – Technik, die begeistert

MEGWARE

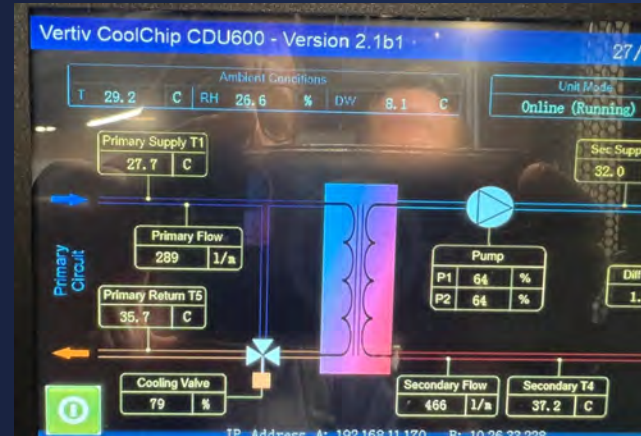
Lenovo

RRZE-Haustechnik

FAU-
Gebäudemanagement

FAU-Verwaltung

Staatl. Bauamt



Weitere BayernKI Systeme



Alex (A100)	TOP500
#184	06/22
#174	11/22
#157	06/23
#186	11/23
#212	06/24
#250	11/24
#292	06/25
#322	11/25
#351	06/26

Alex Cluster (NHR@FAU)

MEGWARE AMD / NVIDIA

656 NVIDIA GPUs

352 A40 (44 nodes)

304 A100 40GB/80GB (38 nodes)

0.5 PB

Lustre on NVME – IB attached

BayernKI access possible



BayernKI/MCML Cluster (LRZ)

Lenovo / NVIDIA

442 NVIDIA GPUs (LRZ 204, MCML 238)

- 51 nodes with 4 H100 94 GB HBM2e
- 26 nodes with 4 A100 80 GB HBM2e
- node-local NVME storage / NDR200 Infiniband

plus further nodes (different types, older generations)

BayernKI share: 120 NVIDIA H100 GPUs

Multi-TB NVMe storage system

- ~ 500TB capacity, IB attached

Multi-PB HDD storage system (LRZ Data Science Storage)

BayernKI – Niederschwelliger Zugang

Stand Zugang/Projektverwaltung



- Zugang zu BayernKI-Ressourcen am [LRZ](#) oder [NHR@FAU](#) via Formular auf Bayern-KI-Webseite



Access

Researchers at Bavarian universities and universities of applied sciences (HAWs) can use the BayernKI resources (CPU & GPU) at LRZ and NHR@FAU free of charge.

Interested? In need of AI resources? You can apply here to get access:

[Apply for Resources](#)

Contact us

Erlangen National High Performance
Computing Center (NHR@FAU)
Martensstraße 1 91058 Erlangen

- hpc-support@fau.de
- hpc.fau.de

Leibniz Supercomputing Centre of the Bavarian
Academy of Sciences and Humanities (LRZ)
Boltzmannstraße 1 85748 Garching

<https://www.ki.bayern.de/wissenschaft>

Registration for BayernKI Access

LRZ and NHR@FAU are providing easy access to basic compute resources (CPU & GPU) within the framework of "BayernKI" to researchers at Bavarian universities. Access within BayernKI will be granted on the level of organizational units (chairs, professorships, labs, etc.) and not project-based.

Please fill in the form below to get access for your organization to the BayernKI infrastructure*

University, e.g. "TH-Deggendorf" *

Organizational Unit (Chair/Lab/Professorship), e.g. "Professur für Robotik" *

Webpage of the Organizational Unit *

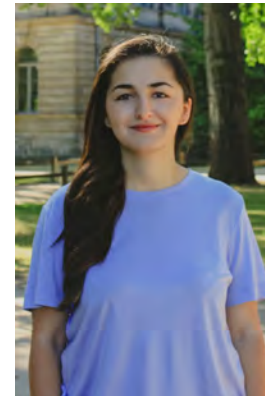
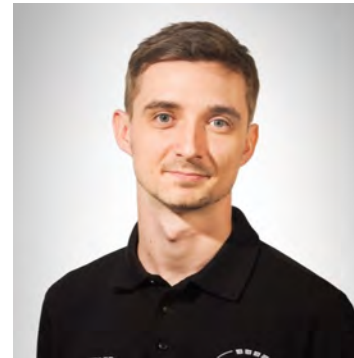
Name and title of the principal investigator (PI) (is responsible for compliance with the export control laws and, thus, should be supervisor of the people using the systems) *

- Projektbeantragung/–verwaltung: Ebene der Lehrstühle / Arbeitsgruppen
- Rollen im Projekt:
 - PI (Principal Investigator)
 - PoC/Master User (Ansprechpartner:in)
 - User (Nutzer:in)
- Wichtig:
 - PI muss Exportkontrollrichtlinien bestätigen!
 - Jedes Projekt erhält Basiskontingent das sofort nutzbar ist
 - Nachgängige Berichtspflicht bei signifikanter Nutzung
- **Ziel: Zugang innerhalb von 3 Tagen**
- **Umfang Basiskontingent: mind. 5.000 GPUh**

- Aufbau eines **gemeinsamen KI-Betreuungs- und Schulungsteams**

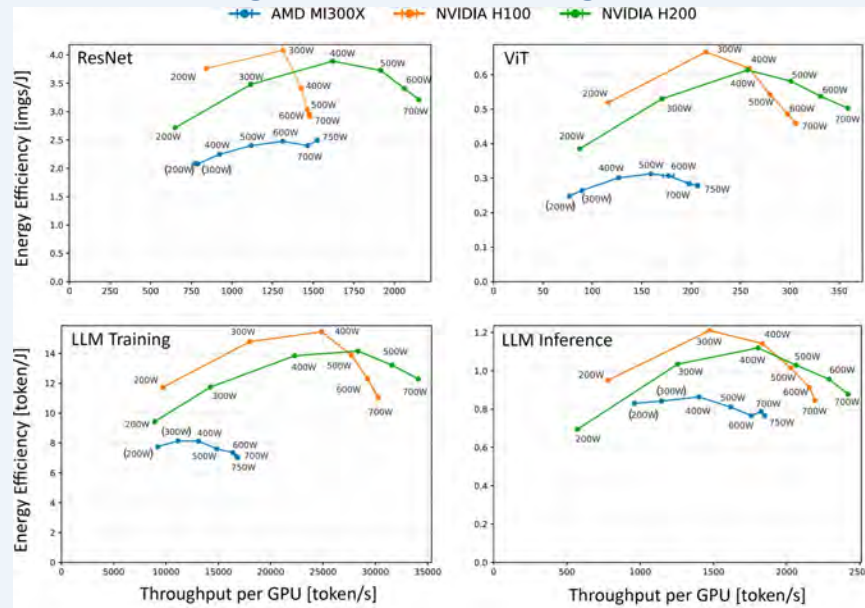
- Zielgruppenorientierte Trainings und Ausbildung
- Mentoring für komplexe, ressourcenintensive Anwendungen

- Lokale Ansprechpartner & Servicemanager

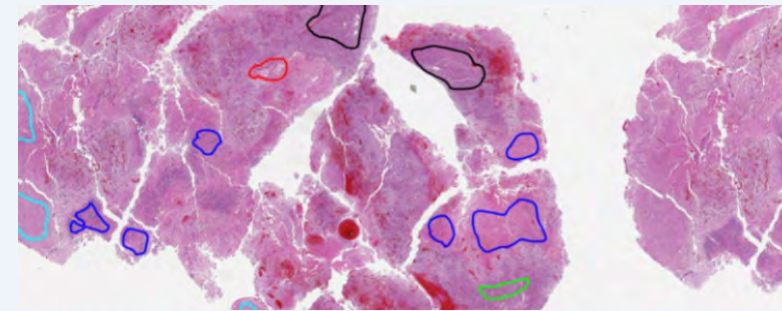
Digitalverbund
Bayern

Martin Mayr (FAU), Mares Barekzai (LRZ), Chang Liu (FAU)
Armine Garibyan (FAU), Benjamin Geißler (LRZ), Lukas Schröder (NHR)

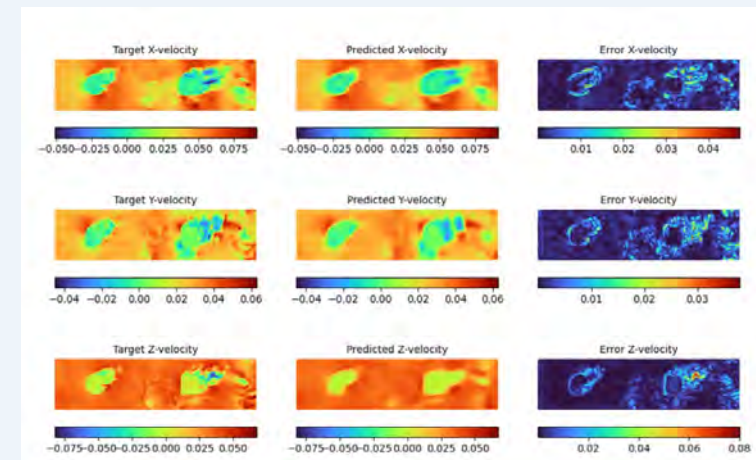
KI Benchmarking und Optimierung



Medizinische Bildverarbeitung



Surrogate Models



LLMs

u. a. Training- und Inference-Benchmarks auf den BayernKI-Systemen



- Q&A BayernKI (LRZ + NHR@FAU)
aktueller Termin: 01.07.2026, 13 Uhr (monatlich)
<https://www.ki-in-bayern.de/support/>



- HPC Café (NHR@FAU) <https://hpc.fau.de/teaching/hpc-cafe/>
nächster Termin: 14.07.2026, 16 Uhr
- AI on HPC in a Nutshell (NHR@FAU)
nächster Termin: 16.07.2026, 16 Uhr
- European AI Hackathon (LRZ)
<https://www.openhackathons.org/s/siteevent/a0CUP00003yKxcX2AS/se000475>

<https://hpc.fau.de/teaching/>

<https://www.lrz.de/en/service-center/alle-lrz-services/beratung/education-and-training>

BayernKI – Ausblick

- **Experimenteller *LLMs as a Service Endpoint***
- **Zielgruppe:** Forschung und Entwicklung an den Univ. & HS
- **OpenAI-kompatible API Key** kann bei am NHR@FAU angefordert werden
 - Front-Ends (z.B. Chat App, OpenCode, ...) müssen selbst eingerichtet werden
 - Keine Speicherung von Fragen und Antworten!
- Aktuelle Open Weights Modelle (können variieren)
 - GPT OSS 120B
 - Qwen3.6 35B
 - Gemma 4 E4B
 - DeepSeek V4 Flash 284B
 - Mistral Medium 3.5 128B
 - ...

<https://hpc.fau.de/request-llm-api-key>

- Nächste Schritte
 - Rahmenvereinbarung → konkrete Bedarfe bitte melden!
 - Erweiterung KI-Hardware LRZ/NHR@FAU Q3 / 2026
 - Abstimmung Digitalverbund
- Kommunikation / Outreach
 - Wir präsentieren BayernKI gerne vor Ort – Seminare / Workshops / ...
 - Feedback!?
- Kontakt: hpc-support@fau.de



<https://www.ki.bayern.de/wissenschaft>



FAQs der letzten BayernKI Vorstellungen



Wie funktioniert der Einkauf eigener Hardware?

Beschaffung über den Rahmenvertrag – die Nodes gehen in das allgemeine System, die Einrichtung erhält dafür zusätzliche Rechenzeit.
Kein Hosting: Die Stromkosten übernimmt das Ministerium.

Wer erhält LLM-API-Keys?

Nur Forschung und Entwicklung – vorläufig nicht für Studierende.

Welche Rechte gelten auf den Clustern?

Keine Adminrechte und keine „hardwarenahe“ Forschung.
Sonderrechte können nach Prüfung auf Anfrage zugeteilt werden.

Fallen Abschlussarbeiten unter die BayernKI-Nutzung?

Ja – Abschlussarbeiten sind abgedeckt.

Gibt es nur GPU-Systeme für die KI-Forschung?

Nein, es stehen auch sehr performante CPU-Cluster zur Verfügung.