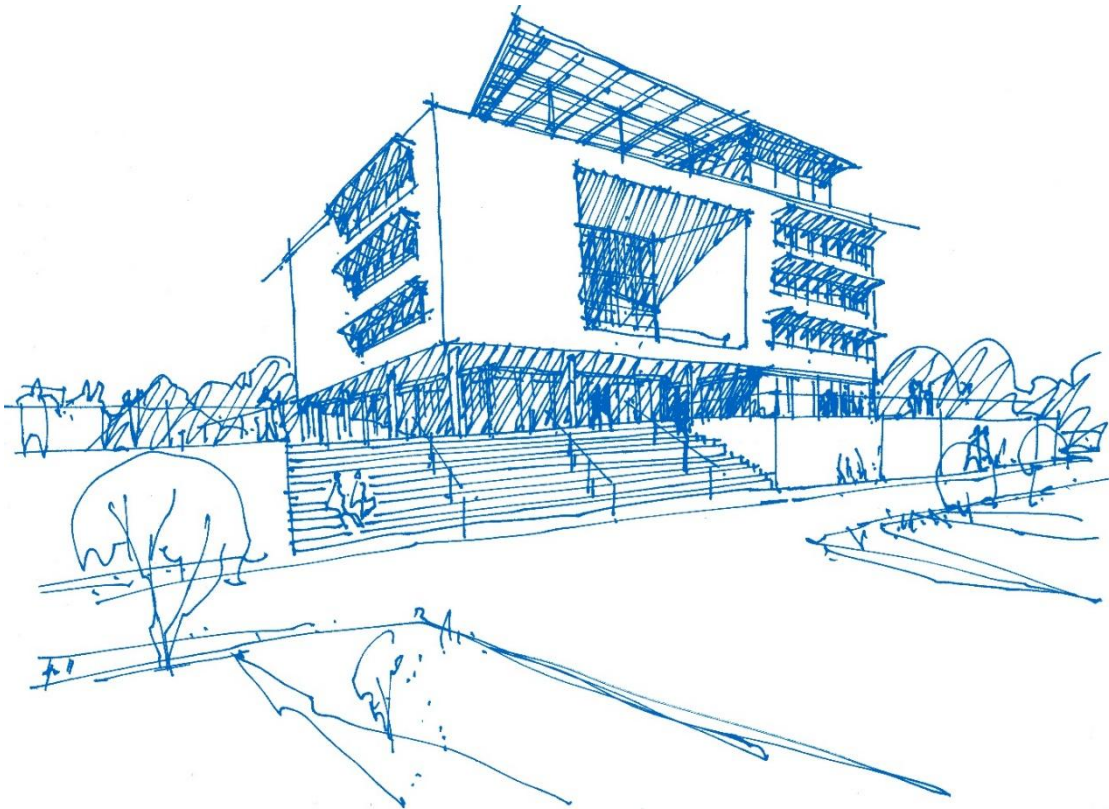


# Zusammenfassung und Empfehlungen

## *Supercomputer, Künstliche Intelligenz und Europas technologische Souveränität*

Workshop 16. und 17. März 2026 am  
**TUM Institute for Advanced Study**



### **TUM Institute for Advanced Study**

TUM School of Computation, Information, Technology

Bayerische Akademie der Wissenschaften (BAW)

European Academy of Sciences and Arts (EASA)

## Vorwort

Internationale Fellows des TUM-IAS forschen mit ihren TUM-Host-Professoren\*Innen und PhD-Studenten\*innen im Bereich „Daten, Deep Learning und Anwendungen“. Hierauf basierend entstand die Idee, den Workshop „Supercomputer, Künstliche Intelligenz und Europas technologische Souveränität“ zu organisieren. Wir stellten uns die Frage, was es für universitäre und staatliche Forschungsinstitutionen in Deutschland und Europa bedeutet, dass große globale Industrie-Player mit enormen finanziellen Möglichkeiten zunehmend eigene Forschung betreiben und die Märkte beherrschen. Auf diesem Hintergrund erschien es uns sinnvoll, das Wissen von Experten aus Universitäten, der Max-Planck-Gesellschaft, der EU und internationalen Unternehmen im Rahmen eines Workshops zu bündeln. Ziel war es, Wege zu einer innovationsoffenen, nachhaltigen und souveränen europäischen Struktur von Großrechnern und KI zu diskutieren und zu beschreiben, die der Forschung restriktionsarm zur Verfügung steht.

Wir danken sehr den TUM-Kollegen Professor Dr. **Klaus Mainzer** (*Präsident European Academy of Sciences and Arts*), Professor Dr. **Arndt Bode** (*Präsident Bayerische Transformations- und Forschungsstiftung und Vizepräsident der Bayerischen Akademie der Wissenschaften*), Professor Dr. **Helmut Krcmar** (*Vizepräsident TUM Campus Heilbronn, 2018 bis 2024*) und Professor Dr. **Hans-Joachim Bungartz** (*Dean, TUM School Computation, Information and Technology*) für Ihre engagierte, wertvolle Unterstützung bei der Konzipierung und Durchführung des Workshops, der Findung der eingeladenen Sprecher und Sprecherinnen und Abfassung vorliegender Zusammenfassung. Erwähnt sei, dass Klaus Mainzer, Arndt Bode und Helmut Krcmar als TUM Emeriti of Excellence Mitglieder der TUM Senior Excellence Faculty sind.

Das Format der am TUM IAS organisierten Veranstaltung mit ausschließlich eingeladenen Spezialist\*innen und deren Kurzvorträgen ermöglichte intensive Diskussionen, z. T. auf Basis vorab eingereichter und verteilter Abstracts. Allen Teilnehmer\*innen sei an dieser Stelle nochmals für ihren wichtigen Input gedankt. Die Ergebnisse der 2-tägigen Veranstaltung sind im vorliegenden Papier, inklusive der Empfehlungen, zusammengefasst.

Prof. Dr. Dr. h.c. Michael Molls  
Emeritus of Excellence  
Speaker TUM Senior Excellence Faculty, 2015-2025  
Director TUM Institute for Advanced Study

Dr. Ulrich Marsch  
Managing Director TUM-IAS

# INHALTSVERZEICHNIS

## EXECUTIVE SUMMARY

### UNIVERSITÄRE FORSCHUNGSPERSPEKTIVEN

#### **JÜRGEN PFEFFER**

TUM School of Social Sciences and Technology: „Daten als Grundlage für neue Forschungsfelder“

#### **CARLA-SOPHIE RIEGER**

TUM/CERN: „Forschung an der Schnittstelle von KI und Hochleistungsrechnen – Perspektiven junger Wissenschaftler:innen“

#### **KLAUS MAINZER**

European Academy of Sciences and Arts (EASA): „Forschung, Lehre und Rechenleistung – quo vadis Europa?“

#### **LEO SCHWINN**

TUM: „Das Munich Data Science Institute der TUM“

### EUROPAS SUPERCOMPUTER-INFRASTRUKTUR

#### **THOMAS LIPPERT**

Jülich Supercomputing Centre (JSC) und Goethe-Universität Frankfurt: „JUPITER for Simulation and AI - Perspectives of the Jülich Supercomputing Centre (JSC)“

#### **DIETER KRANZLMÜLLER**

Leibniz-Rechenzentrum, Garching, Bayerische Akademie der Wissenschaften: „Blue Lion – der neue bayerische Supercomputer und photonische Ansätze“

#### **ERWIN LAURE**

Max Planck Computing and Data Facility: „Zukunft des High-Performance-Computings“

### UNTERNEHMENSSTRATEGIE UND DIGITALE SOUVERÄNITÄT

#### **HELMUT KRCMAR**

TUM: „Unternehmensstrategien und digitale Souveränität“

#### **ROBERT JOZIC**

Dieter Schwarz Digital GmbH: „STACKIT Cloud als europäisches Beispiel für digitale Eigenständigkeit“

#### **SAI NARASIMHAMURTHY**

ParTec AG: „Sustaining Europe’s Leadership in Advanced Computing through Community engagement and co-design“

### DEUTSCHE UND EUROPÄISCHE GROßFORSCHUNG

#### **PHILIPP ULBL**

MPI für Plasmaphysik: „Fusion Research, Supercomputing, and AI“

#### **KILIAN GROSS**

EU-Generaldirektion CONNECT: „Technologische Souveränität und Innovation“

### ENERGIEEFFIZIENZ UND NACHHALTIGKEIT VON SUPERCOMPUTERN

#### **AXEL AUWETER**

MEGWARE GmbH: „Energieoptimierung und Kühlstrategien im HPC-Betrieb“

**IULIA YAMNENKO**

MPI Quantenoptik: „Großrechner und KI in Konfliktzeiten: Resilienz, Potenziale und Partnerschaften am Beispiel der Ukraine“

## **EMPFEHLUNGEN**

## EXECUTIVE SUMMARY

### **Digitale Souveränität ist eine Grundvoraussetzung für die Souveränität Europas**

Europäische Souveränität muss von der europäischen Vision seiner selbstbestimmten Zukunft geleitet werden, auf den Grundfesten von Demokratie, Grundrechten und Rechtsstaatlichkeit.

Souveränität bedeutet nicht Autarkie, sondern das Vorhandensein von Wahlmöglichkeiten, die Handlungsfähigkeit ermöglichen. Reale Wahlmöglichkeiten (Optionen) kosten. Pragmatisch formuliert: „Europa muss sich selbst genug sein können – auch wenn es das nicht will, auch wenn es das nicht immer tut und auch wenn es manchmal teurer wird.“

Wer aufholen will, muss schneller laufen als die Führenden bzw. muss sich in die Lage versetzen, in naher Zukunft schneller laufen zu können. Konkret bedeutet das, dass Ressourcen für digitale Dienste nach Europa geleitet werden müssen; mindestens 30% der IT-Ressourcen sollen in die Förderung und Nutzung souveränitätsstiftender Lösungen fließen. So werden die Ressourcen für die Aufholjagd in Europa aufrechterhalten.

Ein Moonshot-Projekt „European HPC und AI-Chips“ ist erforderlich, möglich und umsetzbar. Dazu braucht es eine klare Zielsetzung, Willen, langen Atem, Ressourcen und Skills. Langer Atem setzt Entschlossenheit voraus.

Es geht nicht nur um Chips und Rechenzentren, sondern um den gesamten TechStack: Die Softwaredimension muss mitgedacht und mitgemacht werden.

HPC-Zentren und Gigafactories für KI<sup>1</sup> müssen sich ergänzen, um hybride Lösungen zu ermöglichen. Simulation ist unabdingbar, kann jedoch durch hybride Ansätze (KI, Quanten, Photonik) verbessert und kostengünstiger gestaltet werden.

Das Mindset muss sich von „Verwalten von Knappheiten“ zu „Verfügbarkeit abundanter Ressourcen“ ändern.

Weg vom „Mein-Dein-Denken“ hin zu „Transparenz der Ressourcen und Open-Innovation-Ökosystemen“. Unsere Herangehensweisen an die KI sollen „brute force“ (USA) mit „high precision“ und „sustainability“ (EUR) verbinden. Die Europäische Lösung: Gewinnen durch mehr Cleverness statt Wettbewerb um noch mehr milliardenteure „Brute Force“.

#### **Hoffnung ist keine Methode: Heute ins Handeln kommen!**

- Digitale Souveränität ist Teil des unternehmerischen und politischen Risikomanagements
- KPI auf Aufsichtsratsebene: Budgetanteil von souveränitätsfördernden Komponenten
- KPI auf der politischen Ebene: systemische Transparenz
- Open Source als Innovationsökosystem fördern
- Mitwirkung bei souveränitätsfördernden Verbänden / Genossenschaften
- Ziel D<sup>2</sup>S<sup>2</sup>: Wir können und sollten Vorbild sein, sowohl im Datenschutz als auch in der digitalen Souveränität!

***Eile ist geboten: Wer aufholen will, muss schneller laufen als die Führenden!***

---

<sup>1</sup> Sehr große Rechenzentren, die speziell für die Entwicklung, das Training und den Betrieb komplexer künstlicher Intelligenz konzipiert sind.

## UNIVERSITÄRE FORSCHUNGSPERSPEKTIVEN

**JÜRGEN PFEFFER**, TUM School of Social Sciences and Technology: „Daten als Grundlage für neue Forschungsfelder“

Der Vortrag behandelt die Erforschung des menschlichen Verhaltens anhand von Daten, insbesondere durch die Analyse von Netzwerken (z. B. Social Media wie TikTok oder Twitter). Dabei geht es nicht nur um die Daten, die Menschen erzeugen, sondern darum, wie sich **Verhalten aus ihnen rekonstruieren und verstehen lässt**.

Ein Schwerpunkt liegt auf **Netzwerkodynamiken** (z. B. darauf, wie sich Themen verbreiten, wie Menschen vernetzt sind und wie sich Strukturen verändern). Methoden wie **statistisches Modellieren** und der Einsatz von **Large Language Models (LLMs)** zur Textanalyse spielen dabei eine wichtige Rolle.

Zentrale Herausforderungen sind:

- Modellierung komplexer Netzwerke (Skalierungsprobleme, z. B. >1000 Knoten schwer handhabbar)
- Technische Skalierung und Rechenleistung
- Fine-Tuning von LLMs (z. B. experimentelle Projekte wie „DemocraGPT“<sup>2</sup>)

Im Vergleich zwischen Deutschland und Österreich wird die **deutsche Dezentralisierung** von Forschung und Unternehmen hervorgehoben. Dem steht ein Trend in der KI gegenüber: die **Zentralisierung von Macht in wenigen Schlüsseltechnologien**.

Die Rolle des Supercomputings wird kritisch betrachtet: Es ist zwar für Spitzenleistungen wichtig, aber nicht unbedingt flächendeckend erforderlich. Stattdessen könnte eine **mittlere Rechenleistung** für viele Anwendungen ausreichen. Gleichzeitig zeigen Datenmengen (z. B. 100 TB pro Tag auf Twitter), dass die Skalierung eine zentrale Herausforderung bleibt.

### Diskussion

Diskutiert wird der Bedarf an **Unterstützungspersonal**, das nicht-technische Disziplinen (z. B. Geistes- und Sozialwissenschaften) bei der Nutzung von Technologien begleitet. Zudem fehlen **IT-affine Fachkräfte** in diesen Bereichen.

Weitere zentrale Punkte:

- **Niederschwelliger Zugang** zu Rechenanlagen und Vermeidung von Leerzeiten sind wichtig
- Die Nachfrage nach Rechenressourcen wird stark steigen
- Nur wenige Projekte benötigen tatsächlich maximale HPC-Leistung
- Große Datenmengen (z. B. Social Media) stellen wertvolle Ressourcen dar
- Es besteht ein Spannungsfeld zwischen **hoher Performance einer Rechenanlage und flexibler Verfügbarkeit**
- Auf struktureller Ebene werden die europäische **Pyramidenstruktur der Nutzer** sowie neue Initiativen wie **KI-Fabriken** erwähnt, die stärker auf Interaktivität setzen.
- Abschließend wird betont, dass Universitäten eine wichtige Rolle spielen sollten, um **Zugang, Nutzung und Kompetenzen zu verbessern**, jedoch ist noch unklar, wer diese Entwicklung konkret vorantreibt.

**CARLA SOPHIE RIEGER** (TUM/CERN): „*Forschung an der Schnittstelle von KI und Hochleistungsrechnen – Perspektiven junger Wissenschaftler:innen*“

---

<sup>2</sup> Entwicklung eines KI-gestützten Gesprächstrainings zur Förderung demokratischen Austausches

Dieser Beitrag beleuchtet diese Themen aus der Perspektive junger Wissenschaftler:innen an der Schnittstelle von Quanten- und Hochleistungsrechnen.

Die Integration von Quantencomputern in HPC-Ökosysteme gilt als eine der zentralen Herausforderungen der kommenden Dekade. Mit der European Quantum Strategy hat die EU einen strategischen Rahmen geschaffen, der diese Entwicklung untermauert – doch wie können Strukturen wie EuroHPC weiterentwickelt werden, um den spezifischen Anforderungen des Quantencomputings (QC) gerecht zu werden? Ebenso gewinnt die Energieeffizienz als Designprinzip künftiger Quantenarchitekturen an Bedeutung – die Quantum Energy Initiative in Grenoble stößt hierzu an, einen tragfähigen Rahmen für die Bewertung von Kosten und Nutzen zu schaffen. Zugleich stellen sich grundlegende Fragen zur Nutzung der klassischen HPC- und QC-Infrastruktur. Es gilt zu erkunden, wie Forschende aus unterschiedlichen Disziplinen dabei unterstützt werden können, diese Systeme durch Parallelisierung, Skalierung und Codeoptimierung effizient einzusetzen. Auch die Entwicklung flexiblerer Zugangsmodelle verdient Aufmerksamkeit, um explorativer QC- und KI-Forschung Raum zu geben – auch wenn sich der konkrete Rechenbedarf im Voraus oft nur schwer abschätzen lässt, gleichzeitig aber eine gute Auslastung der Systeme gewährleistet werden soll.

### **Zentrale Herausforderungen & Fragen**

- Klassische Simulation von Quantensystemen stößt an Grenzen, da sie exponentiell skaliert → limitierte effiziente Simulierbarkeit
- Unklare Grenzen der klassischen Skalierbarkeit vs. Potenziale von Quantencomputing
- Wie lassen sich in Bezug auf die Datenprozessierung HPC und Quantencomputing effizient verknüpfen?
- Wie weit sind wir noch von skalierbaren, fehlertoleranten Quantencomputern entfernt?

### **Technologische Ansätze**

- Hybride Methoden (Kombination aus klassischem HPC und Quantencomputing) als Schlüsselansatz
- Integration von Quantencomputern in HPC-Systeme
- Nutzung von Quantencomputern als Beschleuniger innerhalb klassischer HPC-Infrastrukturen

### **Ressourcen & Nutzung**

- Häufige Diskrepanz zwischen verfügbaren und tatsächlich genutzten Ressourcen in HPC-Workflows
- Ziel der effizienteren Ressourcen-Nutzung (z. B. durch Optimierung, Skalierung)
- Industriekooperationen unterstützen Forschungsvorhaben, ersetzen aber keine offene europäische Infrastruktur

### **Menschen & Zugang**

- Junge Forschende brauchen besseren Zugang, Training und geeignete Infrastruktur

### **Zukunftsperspektiven & Bewertung**

- Fokus könnte sich von „Quantum Advantage“ hin zu „Quantum Energy Advantage“ verschieben (siehe „Grenoble Quantum Energy Initiative“)
- Neben Leistungsfähigkeit werden Kosten und Energieeffizienz auch im Bereich Quantencomputing entscheidend.

## **KLAUS MAINZER**, EASA: „Forschung, Lehre und Rechenleistung – quo vadis Europa?

Die Entwicklung von Hoch- und Höchstleistungsrechnern ist in der europäischen Technik- und Wissenschaftsgeschichte tief verwurzelt. Nach dem 2. Weltkrieg wurde in Deutschland zwar eine Rechnerinfrastruktur ausgebaut, die dem Bedarf von Wissenschaft und forschender Industrie gerecht werden sollte. Die weltweite Führerschaft ging aber z.B. an die USA verloren. Um zukünftig die Souveränität und Unabhängigkeit Europas und Deutschlands politisch, wirtschaftlich und militärisch zu sichern, muss im weltweiten Wettbewerb eine technologische Führungsrolle angestrebt werden. Dazu müssen in Deutschland künftig die besten und leistungsfähigsten Rechnertechnologien zur Verfügung stehen. Der Treiber dieser Entwicklung ist der dramatisch ansteigende Bedarf an Rechenleistung für KI-Infrastrukturen (z.B. KI-Factories).

Rechner und Rechenzentren sollen aber nicht nur leistungsfähig sein, sondern auch energieeffizient und ökologisch nachhaltig. In dieser Forderung kommt ein Leistungs- und Forschungsprofil zum Ausdruck, das den europäischen Werten entspricht und sich von anderen Wettbewerbern abhebt. Dieses Wertebewusstsein sollte im Sinne der „Sustainable Development Goals“ (SDG) der UN bereits in der Ausbildung zukünftiger Entwickler und Anwender zum Ausdruck kommen. Die Europäische Akademie der Wissenschaften und Künste baut derzeit eine internationale und interdisziplinäre KI-Plattform für diese Ziele auf.

### **Hauptpunkte zur europäischen Supercomputing-Strategie**

- Europa muss Überregulierung und Bürokratie vermeiden, um Innovationen nicht zu bremsen.
- Abhängigkeit von Energie- und Technologieimporten soll reduziert und eigene erneuerbare Energie und neue smarte (modulare) nukleare Energie gefördert werden.
- Rüstungsindustrie wird zunehmend relevant für die Entwicklung von Großrechnern.
- Deutschland/Europa muss mit neuen (!) Forschungsideen und Innovationen, die nicht unbedingt kapitalabhängig sind, punkten.
- Deutschland bietet hochwertige und preiswerte Lehre, verliert jedoch viele Absolventen an andere Länder.  
Bei hohen Investitionen ist in einer Demokratie eine gesellschaftliche Einbindung notwendig.
- Wirtschaftlicher, politischer und militärischer Druck erfordert multilaterale Wettbewerbsfähigkeit.

### **Technologische Trends und Chancen**

- Drohnen, KI und neue Atomwaffen treiben den Bedarf an Rechenleistung an.
- Statistische Datenoptimierung im „Machine Learning“ der bisherigen KI muss durch physikalisches Wissen (partielle Differentialgleichungen) erweitert werden, um Lösungen zu verbessern und zu präzisieren (Physical AI).
- Großrechner dienen daher nicht nur der statistischen Optimierung großer Datenmengen mit neuronalen Netzen (Machine Learning) wie in der bisherigen KI, sondern auch der Lösung komplexer mathematischer Gleichungen in z.B. Natur-, Ingenieur- und Wirtschaftswissenschaften.
- Neuromorphe Chips nach dem Vorbild des Gehirns (neuromorphes Rechnen) können den von-Neumann-Flaschenhals<sup>3</sup> vermeiden und den Energieverbrauch senken.
- Memristor<sup>4</sup>-Photonik-Kombinationen im neuromorphen Computing gelten als zukunftssträchtige und energieeffiziente Technologien.
- Photonische Chips eröffnen zudem die Verbindung zum Quantencomputing.

---

<sup>3</sup> Das grundlegende Tempolimit in modernen Computern

<sup>4</sup> Der Memristor ist ein elektronisches Bauelement, eine Schlüsseltechnologie für energieeffiziente KI-Hardware und neuromorphes Rechnen.

- Die Kombination aus statistischer Datenoptimierung und physikalischem Wissen (Physical AI) wird als neuer Innovationsschub gesehen, der sowohl neue KI-Software (automatisches Differenzieren) als auch Superrechner erfordert.

## Diskussion & Empfehlungen

- In Deutschland müssen Investitionsentscheidungen stärker gerechtfertigt werden, während andere Länder schneller Mittel bereitstellen.
- Energieeffizienz (z.B. im neuromorphen Computing und im Quantencomputing) ist ein verständliches und nachhaltiges Thema, das mit dem Green Deal der EU ein typisches europäisches Forschungsprofil darstellt.
- Durchbrüche und neue Ideen in KI- und Superrechnertechnologie hängen nach wie vor von menschlicher Intelligenz und Grundlagenforschung ab, mit deren Förderung Deutschland/Europa auch ohne den Kapital- und Energieeinsatz der Großmächte punkten kann.
- KI benötigt zwar wachsende Rechnerleistung (z.B. KI-Fabriken), sollte aber effizient, erklärbar und kontrollierbar sein, und damit ein typisch europäisches Profil entwickeln.

## LEO SCHWINN, TUM: „Das Munich Data Science Institute MDSI der TUM“

MDSI is essentially a new central hub of TUM for anything related to Data Science and AI. The idea is to facilitate collaboration between people working on AI on a fundamental level, for example, developing new AI methods and researchers in other fields like physics, medicine, management, and others who have a lot of data they want to analyze, where AI could be useful.

The goal of the MDSI is to provide infrastructure, education, and support to connect people from different disciplines to work effectively together.

The MDSI also provides support for data management. For example, by regularly conducting courses that teach specific fundamentals of AI research. But also through personnel, by providing a central point where people can ask these technical questions, so that they do not have to become experts in AI and computer science.

Ideally, we can connect people with diverse skills to solve interdisciplinary problems effectively.

To accomplish this, the MDSI manages a set of computers accessible to its affiliated members.

Similar to LRZ and other compute clusters, this reduces users' workload. Specifically for people from other fields, this can be helpful. However, most of the computing resources used in MDSI-related projects are somewhat decentralized, as most PIs affiliated with the MDSI have their own resources. For example, the GPUs in their lab (LRZ) are resources from Helmholtz Munich, and so on.

To facilitate these research projects, the MDSI also provides seed funds for smaller-scale interdisciplinary projects. Many successful projects have already been conducted, with some receiving awards.

A list of projects can be found on the website.

## Zentrale Punkte zur Nutzung von Rechenressourcen in Deutschland/Europa

- **MDSI** verfügt über eigene Compute-Ressourcen, die Mitglieder nutzen können, um interdisziplinäre Projekte zu starten.
- An der **TUM** sind Doktorand:innen oft nicht geschult, wie sie Rechenzeiten beantragen; das läuft meist über die Professor:innen.
- Für „kleine“ Nutzer:innen ist der Zugang zu HPC-Ressourcen in Europa schwierig; individuelle Forschungszugänge fehlen.

## **Diskussion & Herausforderungen**

- Das Wissenschaftsministerium möchte Rechenressourcen für die Wissenschaft sichern, um eine kommerzielle Nutzung nicht zu bevorzugen.
- Grundlagenforschung verlagert sich zunehmend in Unternehmen, die teilweise über eigene Rechenkapazitäten verfügen.
- Giga-Factories und kommerzielle Anbieter (z. B. Amazon) sind gewinnorientiert und unterscheiden sich von klassischen Forschungsrechenzentren (z. B. LRZ Jülich).
- Föderalismus führt oft zu vielen kleinen, unkoordinierten Anfragen; eine gebündelte Nutzung wäre effizienter.

## **Perspektiven & Empfehlungen**

- HPC sollte primär als Werkzeug für wissenschaftliche Spitzenprojekte genutzt werden; KMU sind bislang kaum eingebunden.
- Kooperation zwischen Wissenschafts- und Wirtschaftsministerium wird als notwendig erachtet, insbesondere für Quantencomputing.
- Neue Geschäftsmodelle könnten Innovationen ermöglichen, während Forschung teilweise in Unternehmen wandert.

# **DISKUSSION UND ZUSAMMENFASSUNG ZU UNIVERSITÄREN FORSCHUNGSPERSPEKTIVEN**

## **Zentrale Themen**

- Europa verfolgt einen anderen Ansatz als die USA, indem Energieeffizienz bei Hochleistungsrechnern (HPC) und in der KI zentral berücksichtigt wird.
- Physical AI und Physical Machine Learning betreffen Software und Hardware; beide Ebenen müssen berücksichtigt werden.
- KI soll menschliche Intelligenz ergänzen, nicht ersetzen, unter Berücksichtigung ethischer Fragestellungen und der Energieeffizienz.

## **Herausforderungen & Nutzung**

- Kommerzielle Angebote (Google Cloud, Amazon Web Services, Azure, Giga-Factories) sind für die Wissenschaft oft zu teuer und nicht optimal für große Modelle geeignet.
- Verteilungsprobleme bei HPC-Ressourcen: Nutzungszeiten sind nicht effizient verteilt, redundante Zeiten sollten besser genutzt werden.
- Der Zugang zu HPC-/KI-Ressourcen muss standardisiert und vereinfacht werden, auch für die Geistes- und Sozialwissenschaften.
- Unterschiedliche Abrechnungs- und Verwaltungsmethoden erschweren die Nutzung für Laien.

## **Ziele**

- Energieeffizienz, Resilienz und Nachhaltigkeit bei HPC- und KI-Systemen in Europa betonen.
- KI soll menschliche Intelligenz ergänzen, nicht ersetzen; dabei ethische Fragen berücksichtigen.
- Physical/Extended ML und AI sind nicht alles – Erklärungssysteme und Denkprozesse sind entscheidend.
- Bessere Strukturen, standardisierte Zugänge und vereinfachte Nutzung für Wissenschaft, KMU und Nicht-Spezialisten.
- Redundanzen in Rechenkapazitäten besser nutzen (z. B. an der TUM) statt knapper Ressourcen über Bürokratie zu verwalten.

- Unterschiedliche Abrechnungsmethoden vereinheitlichen; den Zugang für alle Fachgruppen erleichtern.
- Fokus auf intelligente, flexible Kapazitäten statt auf maximale Größe; spezialisierte vs. universelle Ressourcen abwägen.
- Europäische Angebote sollen unabhängig von Chiplieferungen optimale Arbeitsbedingungen schaffen und Abundanz und Diversität fördern.
- Giga-Factories und US-Angebote sind für die deutsche Wissenschaft derzeit nicht ideal.
- Ziel: HPC- und KI-Systeme effizient, nutzerfreundlich und nachhaltig betreiben, technische und kulturelle Aspekte der Ressourcennutzung berücksichtigen.

### **Strategische Empfehlungen**

- Europa sollte Strukturen schaffen, die unabhängig von Chiplieferungen beste Arbeitsbedingungen ermöglichen.
- Statt weiterer bürokratischer Regeln: Abundanzkultur fördern, Ressourcen transparent und effizient verteilen, Diversität und den Wert der Ressourcen betonen.
- Narrative entwickeln, die technische Komplexität verständlich machen, z. B. zu LLMs, Quantencomputing oder Chip-Herstellung.
- Eine Kultur der Abundanz könnte in Einrichtungen wie der TUM durch Transparenz und die positive Nutzung der Ressourcen etabliert werden.

## **EUROPAS SUPERCOMPUTER - INFRASTRUKTUR**

**THOMAS LIPPERT**, Director Jülich Supercomputing Centre (JSC) und Goethe-Universität Frankfurt: *“JUPITER for Simulation and AI - Perspectives of the Jülich Supercomputing Center (JSC)”*

I would like to give you a brief insight into the planning, construction, and use of the JUPITER exascale computer. JUPITER's name says it all: it stands for Joint Undertaking Pioneer for Innovative and Transformative Exascale Research. JUPITER was planned in a long process dating back to 2009. During this time, not only did new techniques for simulating complex systems develop, but also so-called DEEP Neural Networks (DNN), as represented today by generative pre-trained transformers. So, it is no surprise that JUPITER was designed as a “dual” system from 2020 onwards. JUPITER can handle the largest simulation calculations and AI training tasks. JUPITER will run the largest climate simulations and also develop the first competitive German and European frontier models. JUPITER is Europe's first step into the exascale dimension and frontier models. This comes with many challenges in terms of energy efficiency, cooling, scalability, resilience, and data systems, which will also have to be overcome in the planned European AI Gigafactories.

### **Thesen zur künftigen Entwicklung von KI, HPC und Europas technologischer Souveränität**

**Förderung der Souveränität in Schlüsseltechnologien:** Deutschland und Europa müssen in den wichtigsten Schlüsseltechnologien Unabhängigkeit erlangen, um selbstbestimmte Prioritäten setzen zu können, Abhängigkeiten zu reduzieren und Monopole zu vermeiden. Dazu gehören u.a. Chips, Cloud-Technologien, künstliche Intelligenz, Quantenkommunikation und Quantencomputing, sowie neue Energietechnologien.

**Bedeutung der künstlichen Intelligenz:** KI ist sowohl eine Schlüsseltechnologie *sui generis* im Sinne einer „General Purpose Technology“ als auch ein dualer Enabler für Schlüsseltechnologien wie Robotik, autonomes Fahren und automatisierte Fertigung auf der einen Seite sowie für Schlüsselwissenschaften wie Biologie, Chemie, Materialien, Klima und Physik auf der anderen Seite.

**Wachsender Einsatz souveräner KI: Souveränität in der KI wird in Bereichen wie dem Finanzwesen, den Versicherungen sowie dem Groß- und Einzelhandel zunehmend wichtig.** Unabdingbar sind Autonomie und Souveränität in Bereichen wie der Sicherheit, der Verteidigung und der kritischen Infrastruktur. Krisenresistenz erfordert souveräne Kontrolle über Daten und Infrastrukturen.

**HPC für KI:** Leistungsfähige HPC-Systeme werden dringend benötigt, um in Deutschland Frontier-Modelle für Wirtschaft, Verwaltung, Verteidigung und Gesellschaft, und nicht zuletzt Wissenschaft, souverän trainieren und nutzen zu können. Der Aufbau von AI-Gigafactories in der Größenordnung von 100.000 GPUs und >50 MW Leistung muss bald erfolgen.

**Wertschöpfung durch Tools und Experten:** Hardware ist nur effektiv, wenn sie klug genutzt wird. Daher muss die Schaffung von Entwicklungstools für Modellierung, Qualitätssicherung und Sicherheit gefördert werden, und die Ausbildung von Expertinnen und Experten sowie von Fachkräften in KI muss Priorität erhalten.

**Quelloffene Modelle nutzen:** Für einen schnellen Start mit eigenen Frontier-Modellen sollte man vorhandene Open-source-LLMs und multimodale Modelle unbedingt nutzen, unter der Bedingung, dass auch Daten, Tools und Trainingsverfahren vollständig offen sind. Wir müssen vermeiden, Räder neu zu erfinden und dabei den Start zu verpassen.

**Der Entwicklung nicht nur folgen:** KI entwickelt sich heute rasant in Richtung von Multi-Agenten-Systemen und RAG<sup>5</sup>. Wissenschaftsministerien und große fördernde Forschungsorganisationen müssen Programme schaffen, damit Universitäten und Forschungsinstitute in die Lage kommen, sich an die Spitze solcher Neuentwicklungen zu stellen.

**Marktanteile steigern:** Das vorrangige Ziel Deutschlands muss sein, seinen Marktanteil im Bereich KI auf mehr als 3 % zu erhöhen. Von geschätzten 190 Mrd. € Umsatz in KI weltweit im Jahr 2023: UNCTAD (2025) erwartet eine Steigerung auf fast 5 Bio. im Jahr 2033. Für Deutschland wird hier ein Umsatz von 150 Mrd. für 2033 prognostiziert. Entscheidende Bedeutung haben dabei First-to-Market-Key-Technologies.

### **Zentrale Themen & Entwicklungen**

- Neuromorphes Computing soll KI-Kapazitäten effizient nutzen.
- Bedarf an Supercomputern steigt kontinuierlich, stark getrieben durch KI-Anwendungen.
- Energieeffizienz ist entscheidend; photonisches Rechnen bietet ein hohes Leistungspotenzial bei sehr niedrigem Energieverbrauch.
- Photonische Chips: vollständig in Deutschland entwickelt, energieeffizient, innovative Alternative zu marktüblichen Chips.

### **Aktuelle Herausforderungen**

- Bisherige HPC-Hardware war marktbasierend, entspricht jedoch oft nicht den spezifischen wissenschaftlichen Anforderungen.
- Nvidia dominiert den Markt (16-Bit-Chips, steigende Kosten von bis zu 50.000 USD pro Chip).
- Quanten- und neuromorphes Computing befinden sich noch im frühen Stadium und eignen sich derzeit nur für spezielle Anwendungen.
- Mangel an Standards für Chiplelets sowie an eigenständiger Hardware-Integration.

---

<sup>5</sup> *Retrieval-Augmented Generation*: eine Technik, bei der Sprachmodelle (wie ChatGPT) vor der Antwortgenerierung gezielt externe oder unternehmensinterne Daten abrufen, um präzisere und aktuellere Informationen zu liefern.

## Zukunftsperspektiven & Empfehlungen

- Aufbau gestaffelter HPC-Strukturen: europäische, nationale und kleinere institutionelle Zentren.
- Geringere Präzision (16 Bit) kann situationsabhängig ausreichen.
- Förderung der Eigenständigkeit Europas bei der Chipentwicklung, ohne vollständige Autarkie anzustreben.
- Partnerschaften suchen, Investitionen in die nationale Chipentwicklung; Vorschlag: 30 % des KI-Beschaffungsbudgets.
- Digitale Souveränität als Teil staatlicher bzw. Europäische Souveränität betonen.
- Positive Beispiele: Bayerischer photonischer Chip, energieeffiziente Systeme in Jülich

## Zentrale Punkte zu JUPITER & europäischer HPC/KI-Strategie

- JUPITER-Projekt begann 2009; seitdem Entwicklung neuer Methoden für komplexe Simulationen und Deep Neural Networks (DNN, z. B. generative Transformer).
- Herausforderungen: Energieeffizienz, Kühlung, Skalierbarkeit, Resilienz und Datenmanagement – auch relevant für geplante europäische AI-Gigafactories.

**DIETER KRANZLMÜLLER**, Direktor Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften und Ludwig-Maximilians-Universität München: „*Blue Lion – der neue bayerische Supercomputer und photonische Ansätze*“

**Project Stargate** ist ein US-amerikanisches Großprojekt zum Bau mehrerer KI-Supercomputer mit enormen Investitionen (insgesamt bis zu 500 Mrd. USD, erste Phase ca. 10 Mrd. USD, letzte Phase etwa 115 Mrd. USD) und einer geplanten Leistungsaufnahme von etwa 5 GW. Der Bedarf an Supercomputern wächst stetig, insbesondere aufgrund des zunehmenden Einsatzes von KI.

Ein zentraler Faktor ist dabei die **Energieeffizienz**. Große Hoffnung ruht auf photonischem Rechnen, das deutlich weniger Energie verbraucht als herkömmliche Technologien und zugleich ein hohes Leistungspotenzial bietet. Besonders bemerkenswert: Ein entsprechender photonischer Chip wurde vollständig in Deutschland entwickelt.

Der neue bayerische **Supercomputer Blue Lion** wird ab Anfang 2027 am Leibniz-Rechenzentrum (LRZ) betrieben. Er liefert etwa die 30-fache Rechenleistung des Vorgängers SuperMUC-NG und verknüpft traditionelle Simulationen mit künstlicher Intelligenz:

## Kerndaten und Architektur

- **Leistung und Prozessoren:** Der Supercomputer basiert auf der leistungsstarken NVIDIA Vera Rubin Architektur und nutzt die nächste Generation der HPE-Cray-Technologie.<sup>6</sup>
- **Investition:** Das Projekt wird vom Bund und dem Freistaat Bayern mit insgesamt bis zu 250 Millionen Euro getragen.
- **Kühlung und Energie:** Blue Lion wird zu 100 % direkt flüssig gekühlt. Mit bis zu 40 Grad warmem Wasser, das durch die Racks fließt, kann die Abwärme direkt zur Beheizung von Büros und Gebäuden im Forschungszentrum Garching genutzt werden.
- **Photonische Ansätze und die Zukunft des Computings:** Um den Energiehunger und die physikalischen Grenzen klassischer Rechenleistung zu überwinden, wird am LRZ intensiv an photonischen und optischen Technologien geforscht.

---

<sup>6</sup> Es handelt sich hierbei um eine konvergente HPC/KI-Supercomputing-Infrastruktur, die die höchste CPU- und GPU-Dichte pro Computer-Rack und die höchste Input/Output-Leistung pro Speicher-Rack bietet.

- **Quanten-Integration:** Mit Ansätzen wie dem photonischen Quantencomputing wird daran gearbeitet, klassische HPC-Prozessoren nahtlos mit optischen Quantenrechnern zu verknüpfen, um hybride Workflows in den Naturwissenschaften und der Kryptographie zu ermöglichen.
- Aktuelle Updates, Prototypen von Blue Lion (Blue Cubs) und wissenschaftliche Einsatzgebiete können direkt auf den Seiten des Leibniz-Rechenzentrums (LRZ) verfolgt werden.

Das Leibniz-Rechenzentrum (LRZ) setzt für **Künstliche Intelligenz** primär auf das System **SuperMUC-NG** und nutzt für extreme Datenmengen den **WSE-2-Superchip** (Wafer-Scale Engine) von Cerebras Systems:

- **WSE-2 (Cerebras Superchip):** Der Chip ist so groß wie ein ganzer Wafer. Er vereint 850.000 Rechenkerne auf einer Fläche und ist speziell auf Deep Learning und die schnelle Verarbeitung massiver Datensätze (Big Data) ausgelegt.
- **SuperMUC-NG:** Der derzeitige Haupt-Supercomputer des LRZ in Garching wurde speziell um Grafikprozessoren (GPUs) erweitert, um klassische wissenschaftliche Simulationen mit Methoden der künstlichen Intelligenz zu verknüpfen.
- **Weitere Technologie:** Zusätzlich erforscht und testet das LRZ neuartige Technologien wie **Quantencomputer** und **photonische (mit Licht arbeitende) Rechner**, um KI-Aufgaben noch energieeffizienter zu lösen. Für **photonische Co-Prozessoren (Q.ANT)** optimieren Forschende klassische Algorithmen (wie Matrix-Vektor-Multiplikationen) so, dass sie auf optischen Einheiten laufen, um den Energieverbrauch drastisch zu senken.

Am Leibniz-Rechenzentrum (LRZ) werden KI-Superchips (wie Cerebras WSE-2 oder photonische Co-Prozessoren) hauptsächlich über standardisierte **HPC-Programmierungsumgebungen und Bibliotheken** angesteuert. Die **Programmierung** erfolgt in der Regel hardwarenah in Sprachen wie **C, C++, Python oder Rust** über AI-Frameworks.

Die **Integration und Programmierung** dieser Systeme am LRZ stützen sich auf folgende Kernpunkte:

- **Standard-Sprachen:** Forschende nutzen Programmiersprachen wie **Python, C++** oder **Rust**, um ihre KI-Modelle und Workloads zu schreiben.
- **Bibliotheken:** Für die eigentliche KI-Entwicklung und das Training kommen weit verbreitete Bibliotheken wie **TensorFlow, PyTorch** oder **scikit-learn** zum Einsatz.
- **Optimierung:** Spezielle APIs, Bibliotheken und Treiber des LRZ abstrahieren die Low-Level-Kommunikation und ermöglichen es der Software, den direkten Zugriff auf KI-Beschleuniger (GPUs oder Spezialchips) effizient zu steuern.

#### **Empfehlungen an Wissenschaft und Politik:**

- KI kann HPC nicht ersetzen, aber sehr gut ergänzen, daher interdisziplinäre Forschung zu HPC und KI ausbauen und stärker mit Anwendungsfeldern vernetzen.
- Offene Forschungsinfrastrukturen und Datenplattformen in Europa weiterentwickeln.
- Ausbildung und Qualifizierung von Fachkräften in HPC, KI und Datenwissenschaften stärken.
- Langfristig in HPC- und KI-Infrastrukturen investieren.
- Europäische technologische Souveränität durch eigene Schlüsseltechnologien und Kooperationen stärken.
- Innovationsfreundliche Rahmenbedingungen schaffen und den Transfer von Forschung in Wirtschaft sowie den Zugang für KMU zu HPC- und KI-Ressourcen erleichtern.

**ERWIN LAURE**, Max Planck Computing and Data Facility: „Zukunft des High-Performance-Computings“

Der steigende Bedarf an Hochleistungsrechnen (HPC) wird durch KI und datengetriebene Wissenschaft weiter verstärkt. Große industrielle Systeme wie „Colossus“ zeigen, dass industrielle KI-Infrastruktur zunehmend leistungsfähiger wird und den Markt dominiert, während klassische akademische HPC im Vergleich sehr klein wirken. Damit verschiebt sich auch das Verhältnis zwischen wissenschaftlicher und industrieller Rechenleistung: Während HPC traditionell stark von Forschungseinrichtungen geprägt war, treiben heute große Technologieunternehmen die Entwicklung maßgeblich voran.

Ein zentrales Problem besteht darin, dass der Markt für Hochleistungsrechnen und KI-Hardware stark von wenigen Akteuren dominiert wird. Besonders Nvidia nimmt eine herausragende Stellung ein, da das Unternehmen nicht nur leistungsfähige Grafikprozessoren bereitstellt, sondern auch ein umfangreiches Software-Ökosystem aufgebaut hat. Diese Kombination aus Hardware und Software schafft starke Abhängigkeiten. Viele Anwendungen, Frameworks und KI-Modelle sind eng an Nvidia-Technologien angepasst, was einen Wechsel zu anderen Anbietern schwierig und kostspielig macht. Hinzu kommt, dass spezialisierte Chips und Beschleuniger, etwa für Berechnungen mit reduzierter Genauigkeit wie 16 - (oder weniger)-Bit-Operationen, immer teurer werden. Gerade solche Formate sind für KI-Anwendungen besonders wichtig, weil sie eine hohe Rechenleistung bei geringem Speicher- und Energiebedarf ermöglichen. Die steigenden Kosten und die Marktkonzentration werfen daher die Frage auf, wie langfristig eine bezahlbare und unabhängige Recheninfrastruktur für Wissenschaft und Wirtschaft gesichert werden kann.

Die bisherige Strategie vieler wissenschaftlicher Einrichtungen bestand darin, verfügbare Hardware am Markt zu beschaffen und sie für die jeweiligen Forschungszwecke einzusetzen. Dieses Vorgehen stößt jedoch zunehmend an Grenzen. Industriell entwickelte Hardware orientiert sich vor allem an den Anforderungen großer kommerzieller Märkte, insbesondere an KI-Anwendungen großer Unternehmen. Wissenschaftliche Anforderungen sind jedoch oft vielfältiger: Sie reichen von numerischen Simulationen über Klimamodelle, Materialforschung und Teilchenphysik bis hin zu Bioinformatik und datenintensiven Experimenten. Nicht jede dieser Anwendungen profitiert im gleichen Maße von denselben Chiparchitekturen. Wenn Rechenzentren lediglich das kaufen, was am Markt verfügbar ist, besteht die Gefahr, dass wissenschaftliche Bedürfnisse nur noch nachrangig berücksichtigt werden. Dadurch können Effizienzverluste entstehen, sowohl bei der Rechenleistung als auch beim Energieverbrauch und bei den Kosten.

Zukunftstechnologien wie Quantencomputing oder neuromorphes Rechnen bieten langfristig interessante Perspektiven, sind aber derzeit noch nicht reif genug, um klassische HPC- oder KI-Infrastrukturen in großem Umfang zu ersetzen. Quantencomputer versprechen für bestimmte Problemklassen enorme Vorteile, befinden sich jedoch weiterhin in einer frühen Entwicklungsphase. Es bestehen noch erhebliche Herausforderungen hinsichtlich der Fehlerrate, der Skalierbarkeit, der Stabilität und der praktischen Nutzbarkeit. Auch neuromorphe Systeme, die sich an der Funktionsweise biologischer Nervensysteme orientieren, könnten insbesondere bei energieeffizienten KI-Anwendungen eine Rolle spielen. Dennoch bleiben beide Technologien vorerst Nischenlösungen und können den kurzfristig stark wachsenden Bedarf an Rechenleistung nicht decken. Für die kommenden Jahre bleibt die Weiterentwicklung klassischer und spezialisierter Halbleitertechnologien daher entscheidend.

Vor diesem Hintergrund wird für die Zukunft ein gestaffeltes System von Rechenzentren vorgeschlagen, das europäische, nationale und institutionelle Ebenen miteinander verbindet. Auf europäischer Ebene könnten besonders große und leistungsfähige Systeme bereitgestellt

werden, die für strategisch wichtige Aufgaben, internationale Forschungsprojekte und große KI-Modelle genutzt werden. Nationale Rechenzentren könnten ergänzend spezifische Anforderungen einzelner Länder abdecken und eine breitere wissenschaftliche Nutzung ermöglichen. Institutionelle Rechenzentren an Universitäten und Forschungseinrichtungen würden weiterhin eine wichtige Rolle spielen, da sie näher an den jeweiligen Nutzergruppen sind und flexibler auf lokale Anforderungen reagieren können. Ein solches mehrstufiges Modell könnte dazu beitragen, Rechenressourcen effizienter zu verteilen und zugleich die Versorgungssicherheit zu erhöhen.

Ein zentraler Punkt ist dabei der Aufbau eigener Chip-Entwicklung und zumindest teilweiser Chip-Produktion in Europa. Ziel ist nicht, vollständige Autarkie zu erreichen, da die globalen Lieferketten in der Halbleiterindustrie hochkomplex sind und die internationale Zusammenarbeit weiterhin notwendig ist. Vielmehr geht es darum, kritische Abhängigkeiten zu reduzieren und die digitale Souveränität zu stärken. Europa sollte in der Lage sein, eigene technologische Schwerpunkte zu setzen und Hardware zu entwickeln, die besser auf wissenschaftliche, industrielle und gesellschaftliche Anforderungen abgestimmt ist. Dazu gehören energieeffiziente Architekturen, spezialisierte Beschleuniger für KI und Simulationen sowie offene Software- und Hardware-Ökosysteme, die die Abhängigkeit von einzelnen Herstellern verringern.

Konkret wird angeregt, einen relevanten Anteil der öffentlichen KI-Budgets in die nationale und europäische Chipentwicklung zu investieren. Eine Größenordnung von etwa 30 Prozent der KI-Fördermittel könnte dafür genutzt werden, eigene Kompetenzen in den Bereichen Design, Fertigung, Integration und Softwareunterstützung aufzubauen. Dies wäre nicht nur aus technologischer Sicht wichtig, sondern auch aus wirtschafts- und sicherheitspolitischen Gründen. Wer über eigene Recheninfrastruktur und Chiptechnologien verfügt, kann Innovationen unabhängiger vorantreiben und ist weniger anfällig für Lieferengpässe, Exportbeschränkungen oder geopolitische Konflikte.

Bereits heute gibt es in Deutschland und in Europa vielversprechende Ansätze. Beispiele sind photonische Chips aus Bayern, die Licht statt elektrischer Signale für bestimmte Rechenoperationen nutzen und dadurch perspektivisch hohe Rechengeschwindigkeiten bei geringerem Energieverbrauch ermöglichen könnten. Auch energieeffiziente Ansätze aus Jülich zeigen, dass im HPC-Bereich neue Architekturen entwickelt werden, die besser auf wissenschaftliche Rechenanforderungen zugeschnitten sind. Solche Projekte verdeutlichen, dass Europa über relevantes Know-how verfügt. Entscheidend wird jedoch sein, diese Ansätze nicht nur als Einzelprojekte zu fördern, sondern langfristig in eine koordinierte Strategie einzubetten.

Insgesamt zeigt sich, dass die Zukunft des Hochleistungsrechnens nicht allein durch den Kauf immer größerer Systeme gesichert werden kann. Notwendig ist eine strategische Kombination aus leistungsfähiger Infrastruktur, eigener Technologieentwicklung, internationaler Kooperation und gezielter Förderung. Nur so kann gewährleistet werden, dass Wissenschaft, Wirtschaft und Gesellschaft auch in Zukunft Zugang zu ausreichender, bezahlbarer und souveräner Rechenleistung haben.

## **DISKUSSION UND ZUSAMMENFASSUNG ZU EUROPAS SUPERCOMPUTER-INFRASTRUKTUR**

### **Zentrale Themen**

- Die Supercomputer-Infrastruktur wurde traditionell für die numerische Simulation unterschiedlichster Anwendungen benötigt. In den letzten 10 Jahren sind zusätzlich zu diesen „klassischen“ Supercomputer-Anwendungen sehr hohe Anforderungen an

Rechenkapazitäten in den Bereichen der künstlichen Intelligenz und der Verarbeitung großer Datenmengen entstanden.

- Supercomputer erfordern für Betrieb und Kühlung enorme Energie und belasten damit Haushalte, die Versorgungsinfrastruktur und die Umwelt. Energieeffizienz und Standortauswahl sind daher von großer Bedeutung.
- Die öffentliche, klassische Supercomputer-Infrastruktur in Deutschland und Europa ist grundsätzlich gut organisiert und wird aus EU-, Bundes- und Landeshaushalten finanziert. Für die besonderen Bedarfe der künstlichen Intelligenz besteht derzeit (noch) keine ausreichende Infrastruktur („AI-Gigafactories“).
- Die Koordination der klassischen Infrastruktur erfolgt in Europa durch das Joint Undertaking (JU) Euro-HPC, in Deutschland durch das Gauss Center for Supercomputing (GCS e.V.) für die Bundeshöchstleistungsrechner und durch das NHR (Verein für Nationales Hochleistungsrechnen) für die Landeshöchstleistungsrechner.
- Supercomputer der höchsten Leistungsklasse werden ausschließlich außerhalb Europas entwickelt und produziert. Das gilt sowohl für die Hardware als auch für die Systemsoftware. Der US-Hersteller NVIDIA hat derzeit ein Monopol auf Prozessorbausteine und Systeme der höchsten Leistungsklasse. Europa ist deshalb in diesen Bereichen nicht autark.
- Deutschland hat besondere Kompetenzen in neuartigen Technologien, die künftig Bedeutung für das Supercomputing (klassisch und für KI) haben können: im Bereich energieeffizienter Technologien die Photonik und im Bereich unkonventioneller Rechnerarchitekturen und Programmiermodelle die Quantenrechner.
- Im Bereich der Algorithmen und der Software für Anwendungen besteht dagegen in Deutschland und Europa Kompetenz auf höchstem Niveau.
- Entsprechendes gilt für die Infrastruktur und den Betrieb von Höchstleistungsrechnern.
- Auch im Bereich des Entwurfs, des Betriebs, der Anwendungen und der Nutzung von Cloud-Systemen bestehen in Europa gute Kompetenzen.

### **Herausforderung und Nutzung**

- Im Bereich der Hardware müssen Kompetenzen schrittweise aufgebaut werden. Dabei sollten die in verschiedenen europäischen Ländern vorhandenen Stärken zusammengeführt, gemeinsam genutzt und weiterentwickelt werden.
- Im Bereich Software müssen die Europäer gemeinsam klären, ob ein schrittweiser Umstieg auf Open-Source-Lösungen die Abhängigkeit von den USA (z.B. Microsoft) reduzieren kann
- Europäische Cloud-Lösungen müssen gezielt gefördert und ihre Nutzung propagiert werden

### **Ziele und strategische Empfehlungen**

- Deutsche und europäische Stärken weiter verstärken, Schwächen oder nicht vorhandene Kompetenzen schrittweise kompensieren und beseitigen – auch wenn das sehr teuer wird.
- Systementwicklung für Forschungsinfrastrukturen und Datenplattformen in Europa fördern – für Hardware, Software und Anwendungen.
- Die Konkurrenz zwischen klassischem Supercomputing und Supercomputing für KI gezielt analysieren sowie mögliche Synergien identifizieren und fördern.
- Abgestufte Leistungspyramide etablieren: Gigasysteme – Superrechner - Cloudsysteme
- Autarkie für Europa schrittweise möglichst weit entwickeln, durch gezielte Förderung neuer Technologien und durch Propagierung der Nutzung bereits existierender europäischer Technologien.

- Unkonventionelle bzw. Nischen-Technologien fördern, die teilweise bereits eine besondere europäische Stärke darstellen: Kurz- und mittelfristig im Bereich energieeffizienter Technologien durch Unterstützung der Photonik, längerfristig im Bereich verschiedener Quantentechnologien, von Quantenrechnern und Quantenprogrammierung.

## UNTERNEHMENSSTRATEGIE UND DIGITALE SOUVERÄNITÄT

**HELMUT KRCDMAR**, TUM: „Unternehmensstrategien und digitale Souveränität“

- Zusammenspiel von Abhängigkeit, Souveränität und Autarkie prägt Unternehmensstrategien; Ziel: Wahlmöglichkeiten bewahren.
- Hyperscaler (Cloud-Backbone)<sup>7</sup> haben nur ca. 15 % EU-Anteil; Gigafactories (5 Standorte, je ~100.000 KI-Chips) als europäische Antwort für Industrie, Forschung und Staat.
- Keine vollständige Autarkie möglich; Abhängigkeiten bestehen auf mehreren Ebenen (Lieferanten, Drittparteien).
- Zentraler Konflikt: Cloud Act vs. europäisches Datenrecht; EU-Rahmenwerk enthält keine Regelungen zu Chips.
- Drei Infrastruktur-Strategien:
  1. Hyperscaler nutzen, welche vorhanden sind
  2. EU-Only für maximale digitale Souveränität
  3. Hybrid/Multicloud
- Europa fehlt es an eigenen Rechenzentren und Chips, um digitale Souveränität zu sichern.
- Abhängigkeit von externen Anbietern (z. B. Microsoft) steigert Kosten und Risiken; Alternativen sind dringend nötig.
- Höchstes Risiko: Unterbrechung der IT- und Supercomputing-Infrastruktur.
- Die staatliche Kontrolle von Supercomputern erfordert klare Strategien: mehrere Großrechner, Outsourcing von Workloads, Vielfalt zulassen.
- Investitionen müssen mit Haushaltsmitteln vergleichbar und vertretbar sein.
- Rechtliche Rahmenbedingungen und konservatives Risikomanagement in Deutschland erschweren schnelle Anpassungen.

### **Empfohlene Maßnahmen für Unternehmen**

- Abhängigkeiten aufdecken und managen:
- Politisch: verpflichtende systemische Offenlegung
- Unternehmerisch: dual sourcing, Open Source, zweite Lieferanten für kritische Komponenten
- Souveränität als Board-Kriterium berücksichtigen.
- Workloads klassifizieren und Zielarchitektur festlegen.
- Heute aktiv werden: unternehmerisches und politisches Risikomanagement, um Abhängigkeiten zu minimieren.

### **Unternehmerische & politische Ansätze**

- Risikomanagement als entscheidender Hebel; Plan B und Radial-Flank-Strategien sind notwendig.
- Neue Zulassungen könnten an die Verpflichtung zur Souveränität gebunden werden.
- Europäische Strategie: interoperable Standards, dual sourcing, Open Source, Innovationsökosystem fördern.

---

<sup>7</sup> Ein Cloud-Computing Anbieter, der Rechenleistung, Speicher und Netzwerkinfrastruktur im extremen Maßstab global bereitstellt.

- Verschiedene europäische Komponenten und Ansätze für das Supercomputing kombinieren, um Abhängigkeiten zu reduzieren.

**ROBERT JOZIC**, Dieter Schwarz Digital GmbH: „*STACKIT Cloud als europäisches Beispiel für digitale Eigenständigkeit*“

### **Digitale Souveränität & Eigenständigkeit**

- Viele deutsche Unternehmen entsprechen digital eher dem Stand von Entwicklungsländern.
- Jurisdiktion (z. B. US-Zugriff, US-Clouds) ist ein zentrales Risiko; das heutige System ist zu fragil.
- Souveränität = Wahrung der Entscheidungsfreiheit; der stärkste Hebel für digitale Eigenständigkeit.
- Teilsouveränität (SEAL Level 4)<sup>8</sup> als realistisches Ziel, nicht als vollständige Autarkie.
- Kernbereiche für Souveränität: Clouds, Cybersecurity, Data & AI, Kommunikation, Workspace.
- Ökosysteme aus vertrauenswürdigen Vertragspartnern fördern Souveränität und verbinden Wissenschaft und Wirtschaft.

### **Technologie & Infrastruktur**

- Cloud-Prinzipien: Security by Design, sichere Datenräume, neue Technologien in sicheren Ökosystemen.
- Software ist ein Bottleneck bei Cloud-Anlagen; Open Source erleichtert die Anschlussfähigkeit und verringert die Abhängigkeit von marktbeherrschenden Anbietern.
- US AI Action Plan (Juli 2025) als Referenzpunkt; deutsche Initiativen z. B. Stackit (Schwarz Gruppe) oder IPAI (Heilbronn).

### **Wertschöpfung & Investitionen**

- KI-Wertschöpfung erfordert Fokus; ein hoher Differenzierungsspielraum in jedem Schritt bis zur Anwendung.
- Investitionen in Grundlagenforschung sind wichtig, um Talente zu gewinnen.
- Interaktionen mit europäischen Unternehmen stärken die Souveränität.
- Brain Drain und fehlendes Investment als kritische Defizite.

### **Strategische Empfehlungen**

- Nicht von Null auf volle Souveränität gehen, sondern Schritt für Schritt auf SEAL-4-Niveau hinarbeiten.
- Entschlossenheit und gezielte Investitionsentscheidungen in Cloud, Data & AI sind entscheidend.
- Nutzung von Open Source und interoperablen Standards als Mittel gegen Marktdominanz.

## **DISKUSSION**

### **Bedeutung von Open Source**

- Open Source als Differenzierungsmerkmal gegenüber Hyperscalern.

---

<sup>8</sup> Sovereignty Effectiveness Assurance Level; Wirksamkeit der Souveränität Sicherheitsstufe (SEAL). Die SEAL-Stufe dient als Mindestsicherheitsstufe.

- Ermöglicht Anschlussfähigkeit, niedrighschwelligem Markteintritt und Kooperation statt Konkurrenz.

### Investitionen & Innovation

- Fokus auf Innovation und Zukunftsthemen bei Investitionen ist entscheidend.
- Talentförderung muss integriert werden.
- Langfristige Planung und Risikobereitschaft sind erforderlich; die europäische Wirtschaft muss Cloud-Probleme selbst lösen.

### Wettbewerb & Marktstrategie

- Freies Wirtschaftssystem basiert auf Wettbewerb; Großinvestoren in den USA behindern teilweise den effizienten Betrieb.
- Differenzierungspotenziale erkennen und mithilfe von Software-Stacks in die Anwendung bringen.
- Die Umsetzung scheitert oft an der Bürokratie; die Substanz muss zuerst geschaffen werden.
- Strategien: Differenzieren, alternative Positionen gegenüber Monopolen einnehmen, Zusammenarbeit in Software-Stacks fördern.

### SAI NARASIMHAMURTHY, ParTec AG: „Sustaining Europe’s Leadership in Advanced Computing through Community engagement and co-design“

This talk reflects on ParTec AG's experiences and illustrates how collaboration across research, industry, and public initiatives can drive technological leadership in Europe.

At the center of this is the dynamic Modular System Architecture (dMSA)—an innovative approach to building scalable HPC and AI supercomputers. Originally developed in Germany through close collaboration between ParTec and the Jülich Supercomputing Center (JSC), the dMSA has evolved into the architectural foundation for a number of Europe’s most powerful systems, including the continent’s largest operational HPC and AI supercomputer, culminating in JUPITER, the first Exascale-class system.

The talk highlights how this architecture has been shaped and expanded through a broad European innovation ecosystem. Collaborative research initiatives—including the DEEP and SEA<sup>9</sup> project families—have advanced modular supercomputing concepts and enabled the integration of emerging technologies. Together with the associated software stack, the ParaStation Modulo Software Suite, these architectural principles now also enable the integration of quantum computing systems with classical HPC/AI infrastructures.

Building on these foundations, ParTec continues to apply the lessons learned in new initiatives such as DARE<sup>10</sup>, which aims to establish a RISC-V–based European supercomputing ecosystem<sup>11</sup>, and SEANERGYs<sup>12</sup>, focused on energy-efficient software stacks for future HPC systems. The same collaborative model is now also being applied to

---

<sup>9</sup> DEEP-SEA (“DEEP – Software for Exascale Architectures”) will deliver the programming environment for future European exascale systems, adapting all levels of the software stack to support highly heterogeneous compute and memory configurations

<sup>10</sup> DA/RE ist eine zentrale, deutschlandweit eingesetzte, cloudbasierte IT-Plattform, die die Koordination und Kommunikation der Netz- und Anlagenbetreibenden über alle Spannungsebenen hinweg ermöglicht.

<sup>11</sup> The RISC-V–based European supercomputing ecosystem is a strategic, multi-billion-euro initiative designed to establish **technological sovereignty** and digital autonomy for the EU. By utilizing the open-standard RISC-V **Instruction Set Architecture (ISA)**, Europe is building a complete, domestic hardware and software stack

<sup>12</sup> SEANERGYs aims to create an integrated European software solution that optimizes resource utilization and reduces the energy used for real-world workload mixes.

consult on and co-develop the next generations of AI infrastructure in Europe, working closely with partners and customers from the onset.

Active engagement with the European HPC community—through organizations such as ETP4HPC<sup>13</sup> and the EuroHPC Joint Undertaking's Research and Innovation Advisory Group (RIAG)—has ensured continuous dialogue with stakeholders and policymakers shaping Europe's advanced computing roadmap.

Ultimately, the talk argues that Europe's strength in advanced computing lies in its **community-driven approach**, bringing together industry, research centers, policymakers, and users to co-design the next generation of supercomputing systems.

### Zentrale Themen & Entwicklungen

- Zusammenarbeit zwischen Forschung, Industrie und öffentlichen Initiativen ist entscheidend, um Deutschland und Europa in Advanced Computing & R&I voranzubringen.
- ParTec: Entwicklung von Extrascale-Systemen hin zu souveränen AI-Datenzentren; EU Strategic Advisory beteiligt.
- dMSA (dynamic Modular System Architecture): modulare, dynamische Ressourcenzusammensetzung → verbessert Skalierbarkeit, Leistung und Energieeffizienz; Grundlage für den Supercomputer JUPITER in Deutschland.
- Modulare Architektur ermöglicht die Anpassung an verschiedene Anwendungen, verbessert die Effizienz und bietet Erweiterbarkeit, bringt jedoch hohe Komplexität mit sich.
- Softwareentwicklung erfordert Konsortien; enge Zusammenarbeit mit großen Forschungszentren (auch außerhalb Europas) verbessert die Anpassung an Anwenderbedürfnisse.
- dMSA-Learnings sind auf AI-Datenumgebungen übertragbar → „modulare“ AI-Fabriken in Entwicklung.
- Kontinuität und langfristige Finanzierung sind entscheidend, um Ideenverlust bei Projektwechseln zu vermeiden.
- Community-Engagement, Kollaboration und Co-Design sind essenziell, um nah an aktuellen R&I-Projekten zu bleiben.

### Diskussion & Herausforderungen

- Open Innovation ist nie vollständig offen, da Akteure eigene Anforderungen haben; die Standardisierung ist schwierig.
- Vergleich mit Indien: Regierungsseitiger Anstoß entscheidend, Zusammenarbeit wichtig, da lokale Quantum-Unternehmen fehlen.
- Fehlende Rechenzentren in Europa; mehrere Großrechnerstandorte und digitale Souveränität sind notwendig.
- Abhängigkeit von Hyperscalern → steigende Kosten; Bedarf an Alternativen und „Plan B“.
- Outsourcing von Workloads und Vielzahl an Anbietern nötig; unternehmerisches Risikomanagement und Risikobereitschaft entscheidend (höchstes Risiko: Unterbrechung).
- Staatliche Verantwortung für Supercomputing; langfristige Investitionen und Optimierung entlang des gesamten Stacks erforderlich.
- Open Source nützlich, aber Standardisierung bleibt Problem

---

<sup>13</sup> ETP4HPC is an Industry-led think-tank guiding the development of the European advanced computing ecosystem

## DISKUSSION UND ZUSAMMENFASSUNG UNTERNEHMENSSTRATEGIE UND DIGITALE SOUVERÄNITÄT

### Zentrale Entwicklungen & Konzepte

- Collaborative Research Initiatives (z. B. DEEP- und SEA-Projekte) fördern modulare Supercomputing-Konzepte sowie die Integration neuer Technologien.
- Europas Stärke liegt in der gemeinschaftsgetriebenen Herangehensweise: Industrie, Forschung, Politik und Anwender arbeiten zusammen, um die nächste Supercomputing-Generation zu gestalten.

### Herausforderungen & Empfehlungen

- Bürokratie als zentrales Hindernis; zuerst Substanz schaffen, bevor Verwaltungsprozesse dominieren.
- Talentförderung und Innovationsfokus sind entscheidend.
- Kooperationen und Flexibilität sind notwendig; die europäische Wirtschaft soll als Treiber im Cloud-Markt agieren.

## DEUTSCHE UND EUROPÄISCHE GROßFORSCHUNG

Großrechner und Supercomputer stellen eine Herausforderung für die deutsche und europäische Großforschung dar. Zum Vergleich mit einem anderen Großforschungsprojekt wird im Folgenden die Fusionsforschung betrachtet, dann die Sicht der EU-Generaldirektion CONNECT berücksichtigt.

**PHILIPP ULBL**, MPI für Plasmaphysik: „*Fusion Research, Supercomputing, and AI*“

- **Politischer Kontext:** Fusionsforschung zur langfristigen Energiesicherung ist von zentraler Bedeutung. Der Energiebedarf des Supercomputing unterstreicht die Bedeutung der langfristigen Energiesicherung.
- **Kernfusion & Konzepte:** Was bedeutet Kernfusion?
  - Ziel: Wasserstoff zu Helium zu fusionieren und dabei Energie zu erzeugen.
  - Ein technischer Ansatz ist der Einschluss **des Fusionplasmas** mittels starker Magnetfelder („magnetischer Käfig“)
- **Herausforderung:** Turbulenter Transport in der Plasmarandschicht bestimmt den Wärme(ab)fluss und damit **die Leistungsfähigkeit zukünftiger Kraftwerke**. Daher ist eine genaue Simulation notwendig.
- **Supercomputing & Simulation:**
  - Tools wie **die GENE-Code-Familie** simulieren Plasmaturbulenzen.
  - Neue Entwicklungen: **GENE-X** ist auf Randschichtturbulenzen in komplexen Geometrien wie Tokamaks und Stellaratoren spezialisiert.
  - Spektrale Algorithmen + GPU = **bis zu 500×** höhere Beschleunigung gegenüber traditionellen Methoden.
- **KI-Einsatz:**
  - KI **könnte** zur Simulation von **Plasmaturbulenzen** eingesetzt werden.
  - **Erste Untersuchungen an Subskalen-Turbulenz erreichen eine Beschleunigung von bis zu 700× durch KI-gestützte Algorithmen , angewendet auf vereinfachte Turbulenzmodelle.**
  - Eine effektive Integration von KI kann die Forschung deutlich beschleunigen.

**KILIAN GROSS**, EU-Generaldirektion CONNECT: „*Technologische Souveränität und Innovation*“

CONNECT entwickelt politische Strategien und setzt sie um, um Europa für das digitale Zeitalter zu rüsten. Dazu gehören Investitionen in Forschung und Innovation sowie die Einführung verlässlicher grüner Digitaltechnologien zum Wohle der Wirtschaft und für ein besseres Leben. In Form von Förderungen, Rechtsvorschriften und politischen Initiativen wird dafür gesorgt, dass Europa bei kritischen Digitaltechnologien wie künstlicher Intelligenz, gemeinsamen Datenräumen, Hochleistungsrechnen, 5G, Mikroelektronik, Blockchain oder Quantentechnologie vorangeht und unabhängig bleibt. Europa soll in den Bereichen Datenwirtschaft und Cybersicherheit weltweit führend werden. Es geht um die Förderung des Binnenmarkts, in dem alle Unternehmen zu gleichen Bedingungen investieren und miteinander konkurrieren, Digitaltechnologien, Produkte und Dienstleistungen entwickeln, vermarkten und nutzen können, und in dem Urheber- und Verbraucherrechte geachtet werden. Hochgeschwindigkeitsnetze für alle Unternehmen und Haushalte sollen entwickelt werden, da es ohne Konnektivität keinen digitalen Wandel gibt. Gefordert wird ein innovativer, fairer und nachhaltiger digitaler Wandel der Gesellschaft, der den Menschen in den Mittelpunkt stellt und demokratische Werte und Grundrechte sowie kulturelle Vielfalt achtet. Indem die digitalen Kompetenzen der Menschen in Europa gefördert werden, soll die digitale Kluft überwunden werden.

### Neues Leitmotiv & Ziele

- Ziel ist die Erreichung technologischer Resilienz in Europa, um in Zeiten internationaler Konflikte zu bestehen. Es geht um Schlüsselstellen („Showpoints“), die unabhängig von globalen Lieferketten sind.
- Ziel ist **nicht die vollständige Autarkie**, sondern genügend Stärke, um europäische Interessen wahrzunehmen (Tech-Souveränitätspaket, Cloud-Development, Quantum Act).
- Gefordert ist der Ausbau von KI-Interkultur: Integration von KI in vertikale Wertschöpfungsketten, Entwicklung von Giga-Fabriken, Rechtsrahmen als Basis für einen „KI-Kontinent“.
- Supercomputing ist nur ein Teil der Strategie für schnelle Entwicklung und Wettbewerbsfähigkeit.

### Konkrete Maßnahmen & Infrastruktur

- München gilt als Vorreiter: Die im Februar 2026 von der Deutschen Telekom und Nvidia im Münchner Tucherpark eröffnete "Industrial AI Cloud" ist eine der größten KI-Fabriken Europas. Das Hochleistungs-Rechenzentrum mit rund 10.000 Nvidia-Blackwell-GPUs dient als spezialisierte Infrastruktur für industrielle KI-Anwendungen, um Datensouveränität zu gewährleisten und das Training komplexer Modelle zu ermöglichen.
- So müssen Infrastruktur und Technik in Europa gestärkt und eine Chip-Wertschöpfungskette ausgebaut werden. Leider gibt es derzeit in Europa kaum eigene KI-Chips.
- Analog zum Chip-Act ist ein Quantum-Act geplant, um ein starkes Quantum-Ökosystem mit Start-ups zu sichern und die Abwanderung europäischer Unternehmen in die USA zu verhindern.

### Diskussion und Fragen

- JUPITER & Hardware-Initiative ist beispielfrei:
- Vor der Einweihung gab es Diskussionen darüber, ob Europa zu spät sei.
- Antwort: Die Kapitalintensität ist hoch, aber unverzichtbar. Ziel darf nicht nur die Anwendung sein, sondern auch der Aufbau eigener Kompetenzen, um nicht zum Second-Hand-Kontinent der Großmächte zu werden.
- Chip-Entwicklung & KI-Fabriken
- Europas Unabhängigkeit muss durch die gemeinsamen strategischen Interessen der Mitgliedstaaten gestärkt werden.

- Der Aufbau konkurrenzfähiger Fabriken (Plan Chip Act 2) ist notwendig. Wie das Beispiel Japan zeigt, zieht ein etabliertes Ökosystem Nachahmer an.
1. Unternehmenslandschaft & Skalierungsprobleme
    - In Europa fehlen häufig leistungsfähige Scale-ups, weil der Aufwand für Unternehmen zu hoch ist, insbesondere im Vergleich zu den USA.
    - Eine kurzfristige Lösung ist die vertikale Wertschöpfung. Langfristig bedarf es allerdings einer umfassenderen Strategie.
    - Risikokapital ist in Europa leider schwer mobilisierbar. Ursache sind häufig die gesetzlichen Rahmenbedingungen in Europa mit zu wenigen Anreizen (im Unterschied zu den USA).
  2. Giga-Factories & Finanzierung
    - Die Kritik lautet: Die Finanzierung ginge meist an „Monopolisten“. Andererseits muss der europäische Anteil sichergestellt werden.
    - Buy-European-Initiative: Ohne Konzentration gibt es weder Effizienz noch Wettbewerbsfähigkeit. Die bundesdeutsche und EU-Gießkanne ist die größte Gefahr für die internationale Wettbewerbsfähigkeit.
    - Netzwerkgedanke und Zusammenarbeit in Europa sollten zwar Markt und Mehrwert schaffen, aber ohne Konzentration auf Effizienzgesichtspunkte wird es nicht gehen.
  3. Supercomputing & Souveränität
    - Europa kann Supercomputing-Unabhängigkeit nur durch Kooperation erreichen, da die Kapitalanforderungen enorm hoch sind.
    - Giga-Factories sind ein Mittel, um den europäischen Markt zu etablieren und technologische Souveränität zu sichern.

## DISKUSSION UND ZUSAMMENFASSUNG DEUTSCHE UND EUROPÄISCHE GROßFORSCHUNG

### Fokus der Forschung

- Früher war die Forschung stärker auf den Kern des Plasmas fokussiert. Heute erkennt man, dass auch Randbereiche (Randschichtturbulenzen) für das Gesamtverhalten entscheidend sind.
- Die Physik der **Plasmaturbulenzen ist ein komplexes Multi-Skalen-, Multi-Physik-Problem mit vielfältigen Phänomenen und Wechselwirkungen.**

### Start-ups & Innovation

- Die europäische Fusionsforschung wurde lange Zeit von ITER (Tokamak) dominiert. Mit 7 Partnern ist das Unternehmen zwar international, aber (zu) groß, teuer und zeitlich verzögert.
- Demgegenüber können Start-ups neue Konzepte verfolgen, bestehende Lösungen optimieren und innovative Ansätze wie High-Temperature-Superconductors (HTS) einsetzen.
- Beispiele: Commonwealth Fusion (USA) nutzt HTS-Technologie. Proxima Fusion (Deutschland) verfolgt das Stellarator-Konzept.

### Perspektive

- Weg der Forschung ist entscheidend („Der Weg ist das Ziel“).
- Optimismus fördert vielfältige Ansätze und neue Akteure, die Probleme adressieren und Innovationen vorantreiben.
- Die EU-Generaldirektion entwickelt politische Strategien, um Europa digital zu rüsten. Hierzu gehören Investitionen in Forschung und Innovation sowie die Einführung grüner Digitaltechnologien. Ziel ist die technologische Resilienz Europas, nicht die

vollständige Autarkie, jedoch genügend Stärke, um europäische Interessen wahrnehmen zu können.

## **ENERGIEEFFIZIENZ UND NACHHALTIGKEIT VON SUPERCOMPUTERN**

Energieeffizienz und Nachhaltigkeit sind Forderungen, die zu Markenzeichen von Europa im Zeichen von Green Deal (EU) und SDGs (UN) werden sollten. Zunächst wird die Strategie eines mittelständischen deutschen Unternehmens vorgestellt. In Zeiten kriegerischer Konflikte ergeben sich neue Forderungen, die am Beispiel der Ukraine betrachtet werden:

**AXEL AUWETER**, MEGWARE GmbH: „*Energieoptimierung und Kühlstrategien im HPC-Betrieb*“

Die MEGWARE Computer Vertrieb und Service GmbH hat sich in ihrer über 35-jährigen Firmengeschichte als einer der europaweit führenden Supercomputing-Spezialisten etabliert. MEGWARE entwickelt, konzipiert, produziert und vermarktet Supercomputer der neuesten Generation, von denen sich zahlreiche in der TOP 500-Liste der leistungsfähigsten HPC-Systeme der Welt wiederfinden.

### **Unternehmen & Marktposition**

- MEGWARE GmbH (seit 35 Jahren) ist ein HPC-Systemanbieter und gehört zu den Top 10 der Großrechner-Anlagenbauer.
- Positionierung als SME: Gesundes Wachstum, Anschluss an das bestehende Technologiesystem sowie die Entwicklung eigener Lösungen zeichnen die Positionierung der MEGWARE GmbH als SME aus.

### **Technologische Kompetenzen**

- Seit über 15 Jahren verfügt die Firma über Erfahrung mit Wasserkühlung, passend zu den hohen Anforderungen an KI-Training.
- Zentral ist die Entwicklung eigener Server- und Modul-Rechenzentrums-Konzepte mit z. B. Schnellen Aufbauzeiten von Kühlmodulen innerhalb einer Woche.
- Die Firma verfügt über interdisziplinäre Kompetenzen in den Bereichen IT, Elektro, Sanitär sowie HPC-Softwareunterstützung.

### **Lieferketten & Produktion**

- Multinationale Zulieferung mit Komponenten aus China, Taiwan und Deutschland liegt vor.
- Eigene Elektronikbaugruppen werden entworfen und produziert, Netzwerkkarten werden zugekauft.
- Allerdings ist Second Sourcing für ein KMU wie die MEGWARE GmbH kaum abbildbar. Lieferanten-Audits nach EU-Regeln sind schwierig umzusetzen.

### **Herausforderungen & Trends**

- Die Schnelllebigkeit des Marktes und die neuen Anforderungen durch den KI-Hype stellen für die Firma eine große Herausforderung dar.
- Eine enge Zusammenarbeit mit Partnern (z. B. bei Kälteanlagen) ist erforderlich.
- Die teilweise Fertigung von Komponenten ist zwar lokal möglich, aber komplexe Platinenkomponenten bleiben kritisch

### **Diskussion – Lieferketten & Second Sourcing im HPC-Bereich**

- Es besteht eine globale Abhängigkeit, da viele Komponenten weiterhin aus internationalen Lieferketten stammen. Das vorhandene Know-how ist bereits in der Firma verankert.
- Second Sourcing ist für ein KMU praktisch nicht umsetzbar: Es gibt zu viele verschiedene Komponenten, die einzeln geprüft oder selbst entwickelt werden müssten.
- Der Aufbau eines vollständig lokalen Ökosystems ist derzeit schwierig, da zentrale Designarbeiten (z. B. für Dell, Novum) in Taiwan entstehen.
- In Europa werden, anders als in den USA, bisher immer noch viele Komponenten global beschafft. Kern-Designs konzentrieren sich in Taiwan.

### **IULIIA YAMNENKO**, MPI Quantenoptik, „*Großrechner und KI in Konfliktzeiten: Resilienz, Potenziale und Partnerschaften am Beispiel der Ukraine*“

Die Ukraine entwickelt sich trotz der anhaltenden russischen Invasion zu einem bemerkenswerten Zentrum für künstliche Intelligenz (KI) und High-Performance Computing (HPC), wobei der Fokus stark auf Verteidigungstechnologie (DefenseTech), Sicherheit und der Integration in europäische Strukturen liegt. Zu den aktuellen Entwicklungen zählen (Stand 2025/2026):

#### **1. Strategie und KI-Entwicklung**

- Die Ukraine strebt bis 2030 eine Position unter den drei führenden Ländern bei der KI-Entwicklung und -Integration im öffentlichen Sektor an.
- Dazu hat die Ukraine den EU-AI-Act übernommen und nimmt an Treffen der European Artificial Intelligence Boards teil, um die KI-Standards an die EU-Standards anzupassen.
- Der staatlich geförderte Verteidigungscluster Brave1 ist zentral für die Entwicklung von KI-Lösungen im Militärbereich und arbeitet eng mit internationalen Partnern wie Palantir zusammen.

#### **2. KI im Verteidigungs- und Sicherheitssektor (DefenseTech)**

- KI wird intensiv für die Verarbeitung von Drohnendaten (FPV-Drohnen), die Zielerkennung und die Schwarmintelligenz eingesetzt.
- Systeme wie „Griselda“ verarbeiten riesige Datenmengen aus Drohnen- und Satellitenaufnahmen in Sekunden, um feindliche Stellungen zu identifizieren.
- Im Rahmen der Palantir-Partnerschaft werden sensible Militärdaten zum Training von KI-Modellen genutzt, um die Analyse des Gefechtsfelds zu verbessern.

#### **3. HPC und Infrastruktur-Investitionen**

- Führende ukrainische Cloud-Anbieter investieren über 150 Millionen Hryvnia (ca. 3,75 Mio. USD) in modernste Hardware, darunter NVIDIA H200 Tensor Core GPUs, um leistungsfähige KI-Modelle (LLMs) lokal trainieren zu können.
- Die Ukraine beteiligt sich aktiv an EuroHPC-Projekten (European High Performance Computing Joint Undertaking) und arbeitet an der Integration in europäische AI-Factories.
- KI wird verstärkt zum Schutz kritischer Infrastrukturen vor Cyberangriffen eingesetzt.

#### **4. Anwendungsbereiche außerhalb des Militärs**

- In GovTech wird KI zur Beschleunigung von Verwaltungsprozessen und zur Übersetzung von Rechtsakten im Zuge des EU-Beitritts eingesetzt.
- Startups nutzen KI zur Optimierung von Energienetzen und zum Umweltmonitoring.

**Die Infrastruktur muss vor physischen Angriffen geschützt werden:**

Es besteht ein hoher Bedarf an spezialisierten Fachkräften, um die ambitionierten Ziele trotz des Krieges zu erreichen. Die KI-Transformation in der Ukraine erfordert

- exponentielles Wachstum digitaler Daten in Mathematik, Physik, Chemie, Biologie
- Energieaufwand von KI als zentrale Herausforderung
- Konzentration globaler KI-Ressourcen: USA (~300 Mrd. USD), China (~200–250 Mrd. USD), EU (~100–120 Mrd. USD).
- Variation nationaler Prioritäten: Indien (Landwirtschaft), EU (ethisch & vertrauenswürdig), USA (Verteidigungstechnologien).
- Ein Fallbeispiel ist Israel, das als kleine Nation, große Wirkung in Cybersicherheit, Verteidigung und KI-Start-ups (z. B. Übernahmen durch Google, ServiceNow) erzielt.
- Demgegenüber ist in der Ukraine
  - Der Anteil an globalem HPC/AI-Ökosystem gering (0,5 % vs. 30–40 % USA/China).
  - Die Anzahl von IT-Spezialisten ist relativ hoch pro Mio. Einwohner.
  - Eine einzigartige Datenressource durch Kriegszeit & digitale Transformation: Human Capital, Verteidigungsdaten, Social Media, Militärmedizin, Informationssicherheit.
- **Fazit:** Ukraine kann nicht primär über Rechenleistung konkurrieren, aber durch Daten & Expertise in Kombination mit globaler Infrastruktur einen wertvollen Beitrag zur Zukunft von KI und HPC leisten.

#### **Diskussion – IT-Lernen & Abhängigkeiten in der Ukraine**

- Basis der IT-Kultur:
  - Universitäten integrieren alle relevanten IT-Komponenten in ihre Curricula.
  - Kriegsbedingungen haben die Entwicklung beschleunigt, insbesondere durch Priorität auf militärische Anwendungen.
  - Hohe Vielfalt an Spezifikationen (z. B. Drohnen) fördert Anpassungsfähigkeit und Lernkurven.
- Abhängigkeit von externen Systemen:
  - Problematische Abhängigkeit von Elon-Musk-Systemen (vermutlich Starlink) wurde größtenteils gelöst, teils durch Abschaltung des Systems.
- Offene Fragen:
  - Welche weiteren Maßnahmen oder Ressourcen sind nötig, um IT-Resilienz und Autarkie weiter zu stärken?

## EMPFEHLUNGEN

- Europas Souveränität muss von der Vision seiner selbstbestimmten Zukunft auch für Großrechner und KI geleitet sein.
- Statistische Optimierung großer Datenmengen (heutige KI des Machine Learning mit Superrechnern) muss durch physikalisches Wissen erweitert werden, um Lösungen zu verbessern und zu präzisieren (Physical AI).
- Die Europäische Lösung: Gewinnen mit „high precision“ (Physical AI) und „sustainability“ (Green Deal), statt mit „brute force“ (USA) und „unbegrenzter“ Energie und Rechenpower!
- Es geht nicht nur um Chips und Rechenzentren, sondern um den gesamten TechStack: Die Softwaredimension muss mitgedacht und mitgemacht werden.
- Folgende Beispiele sind erste Schritte im europäischen Supercomputing und KI: JUPITER/Jülich (Erster Exascale-Rechner Europas, weltweit Platz 4), JEDI/Jülich (Platz 1 auf der Liste der energieeffizientesten Supercomputer weltweit im Green500-Ranking), KI-Fabrik München (Telekom/Siemens/Nvidia) mit 10 000 Nvidia-Blackwell-GPUs und einer Rechenleistung von 0,5 ExaFLOPS.
- Neuromorphe Chips nach dem Vorbild des Gehirns (neuromorphes Rechnen) können den von-Neumann-Flaschenhals vermeiden und den Energieverbrauch senken.
- Quantencomputing überwindet zwar die Berechenbarkeitsgrenzen klassischer Computer, erfordert jedoch zunächst klassische Hochleistungsrechner als Simulatoren.
- Energieeffizienz (z.B. im neuromorphen Computing und im Quantencomputing) ist ein nachhaltiges Thema, das im Rahmen des Green Deals der EU ein typisches europäisches Forschungsprofil darstellt.
- Europa muss Überregulierung und Bürokratie vermeiden, um Innovationen im weltweiten Wettbewerb nicht auszubremsen.
- Europa muss in Grundlagenforschung investieren, um Talente zu gewinnen.

## Teilnehmer\*innen und Vortragende

**Axel Auweter**, Chief Technology Officer at Megaware Computer GmbH

**Professor em. Dr. Arndt Bode**, Lehrstuhl für Rechnertechnik und Rechnerorganisation TUM, Präsident Bayerische Transformations- und Forschungstiftung, Vizepräsident der Bayerischen Akademie der Wissenschaften

**Kilian Gross**, Director for Enabling and Emerging Technologies, DG CNECT, Directorate C at the European Commission

**Robert Jozic**, Geschäftsleiter der Schwarz Digital GmbH & Co. KG

**Professor Dr. Dieter Kranzlmüller**, Lehrstuhl für Kommunikationssysteme und Systemprogrammierung LMU, Direktor des Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften

**Professor em. Dr. Helmut Krcmar**, Lehrstuhl für Wirtschaftsinformatik und Geschäftsprozessmanagement TUM, Vizepräsident TUM Campus Heilbronn, 2018 - 2024

**Professor Dr. Erwin Laure**, Lehrstuhl für Rechnerarchitektur & Parallele Systeme TUM, Director of the Max Planck Computing and Data Facility (Garching) Max-Planck-Gesellschaft

**Professor Dr. Thomas Lippert**, Lehrstuhl für Modulares Supercomputing und Quantencomputing Universität Frankfurt am Main, Direktor des Jülich Supercomputing Centre (JSC) am Forschungszentrum Jülich.

**Professor em. Dr. Klaus Mainzer**, Lehrstuhl Philosophie und Wissenschaftstheorie TUM, Präsident European Academy for Sciences and Art (EASA), Salzburg

**Professor em. Dr. Michael Molls**, Lehrstuhl Strahlentherapie und Radioonkologie TUM, Director Institute for Advanced Study TUM

**Sai Narasimhamurthy**, Senior Technologist, ParTec AG

**Professor Dr. Jürgen Pfeffer**, Lehrstuhl für Computational Social Science, School of Social Sciences and Technology TUM

**Carla-Sophie Rieger**, PhD-Candidate, TUM und CERN

**Dr. Ing. Leo Schwinn**, Munich Data Science Institute TUM

**Dr. Philipp Ulbl**, Group Tokamak Theory, Max-Planck-Institut für Plasmaphysik

**Professorin Dr. Iuliia Yamnenko**, Projektmanagement, Max-Planck-Institut für Quantenoptik

## Abkürzungen

DNN	= Deep Neural Networks
D <sup>2</sup> S <sup>2</sup>	= Design to Services and Systems
EASA	= European Academy of Sciences and Arts
GPU	= Graphics Processor Unit
HPC	= High Performance Computing
HPE	= Hewlett Packard Enterprise
KMU	= Kleine und mittelgroße Unternehmen
KPI	= Key Performance Indicators
LLM	= Large Language Model
LRZ	= Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften
MDSI	= Munich Data Science Institute, TU München
QC	= Quantencomputing
RAM	= Random Access Memory
SEAL	= Sovereignty Effectiveness Assurance Level
SME	= Small and medium-sized enterprises
TB	= Terybite
UNCTAD	= United Nations Conference on Trade and Development

Bearbeitung:

Prof. Dr. Michael Molls  
Dr. Ulrich Marsch  
Prof. Dr. Klaus Mainzer  
Prof. Dr. Arndt Bode

Herausgeber:

Technische Universität München  
Institute for Advanced Study  
Lichtenbergstr. 2 a  
85748 Garching  
[www.ias.tum.de](http://www.ias.tum.de)

Juni 2026