

Neural Spacetime Lab in Mannheim.

Die nächste Dimension
der Datenerfassung für
die KI-Forschung.



A²IR

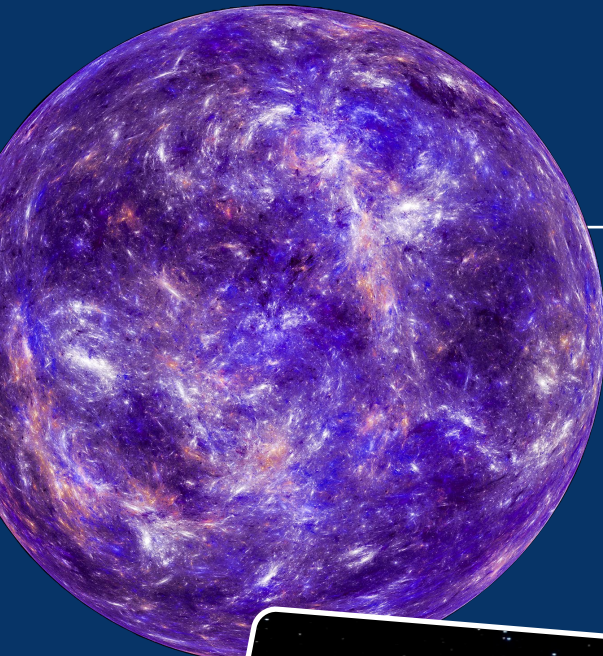
Hochschule Mannheim

Das Institute for Applied Artificial Intelligence and Robotics (A²IR) an der Hochschule Mannheim hatte die wichtige Aufgabe, das Neural Spacetime Lab zu errichten.

Dieses Labor ist als interdisziplinäre Forschungsplattform gedacht und verbindet die Bereiche Computer Vision, Robotik, künstliche Intelligenz und menschliche Interaktion. Sie ermöglicht die präzise Erfassung und Analyse von menschlichen Bewegungen in Echtzeit, was für Anwendungen in der medizinischen Forschung und Robotik entscheidend ist.

Mission: Die Installation fortschrittlicher künstlicher Intelligenz, um die neueste Technologie in intelligenten Kameras zu nutzen, die perfekt aufeinander abgestimmt sind.

Hierbei stieß das Labor auf mehrere Herausforderungen:



- Synchronisierte Erfassung menschlicher Interaktionen: Hochauflösende und echtzeitfähige Erfassung aus allen Raumrichtungen.

- Effiziente Datenerfassung und -verarbeitung: Erkennung, Analyse und Vorhersage komplexer Muster menschlicher Interaktionen.

- Technologische Anforderungen: Intelligente Industriekameras, leistungsfähige Datenaufnahmeserver und fortschrittliche KI-Algorithmen.

Die größte Herausforderung bestand jedoch darin, räumliche Datensätze um die Dimension der Zeit zu erweitern.

Gemeinsam Zukunft formen.



Bundesministerium
für Bildung
und Forschung

DELTA Computer Products GmbH

DELTA, ein führendes Unternehmen in der KI- und Technologiebranche mit Hauptsitz in Reinbek, steht seit über 35 Jahren für innovative Hochleistungsrechenlösungen.

Standort: Reinbek bei Hamburg

Branche: Spezialist für KI-Hochleistungscomputing (HPC) und maßgeschneiderte Storage- und Clusterlösungen

Team: Erfahrene KI- HPC- und Cluster-Experten beraten Sie gern

Zertifizierungen: NVIDIA Elite Partner

Website: www.delta.de

E-Mail: info@delta.de



A²IR

Hochschule Mannheim

Hochschule Mannheim – Institute for Applied Artificial Intelligence and Robotics (A²IR)

Das Neural Spacetime Lab an der Hochschule Mannheim, Teil des Institute for Applied Artificial Intelligence and Robotics (A²IR), ist ein führendes Forschungsinstitut für innovative KI-Assistenzsysteme in den Bereichen Medizin, Logistik und nachhaltige Technologien.

Standort: Mannheim

Kategorie: Forschungseinrichtung

Träger: Land Baden-Württemberg

Forschungsschwerpunkte:
Erforschung von KI-Assistenzsysteme
in der Medizin und Industrie

Team: KI-Forscher und
Medizintechniker / Mediziner

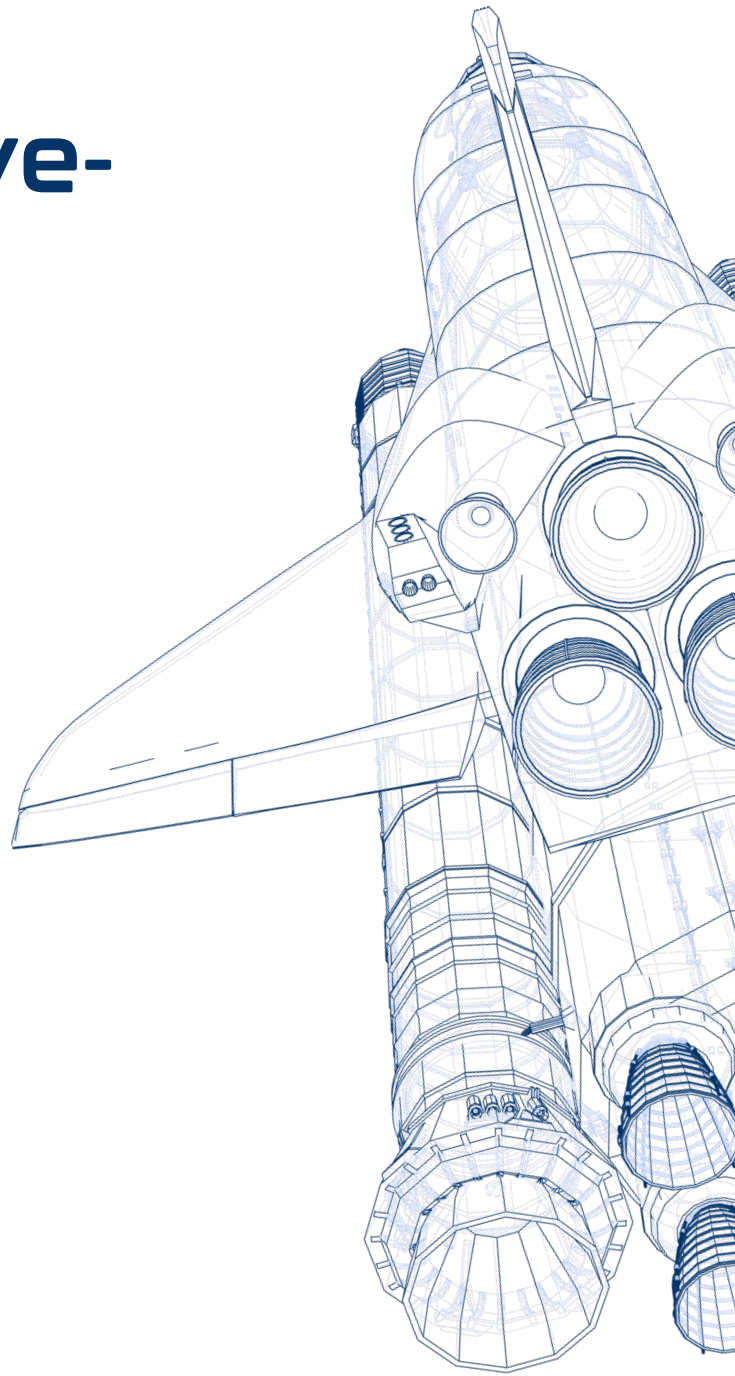
Website: www.a2ir.de

E-Mail: m.vetter@hs-mannheim.de

Die Hyperdrive-Lösung aus Reinbek bei Hamburg

Delta, mit seiner langjährigen Expertise und dem breiten Angebot an NVIDIA KI-Lösungen, konnte einen entscheidenden Beitrag leisten. Für zuverlässige KI-Modelle sind hochwertige und umfangreiche Daten unerlässlich. Schlechte oder unvollständige Daten können die Leistung von KI-Modellen stark beeinträchtigen, was zu ungenauen oder unzuverlässigen Ergebnissen führt. Die Daten müssen sorgfältig aufbereitet und gekennzeichnet werden, was oft viel Zeit und Fachwissen erfordert.

Mit einer Produktpalette von einfachen GPU-Servern bis hin zu hochkomplexen NVIDIA DGX SuperPOD™ Cluster Architekturen bot Delta eine maßgeschneiderte Lösung für die Anforderungen des Neural Spacetime Labs.



Bildquelle: AdobeStock_465702194

„Wir freuen uns, dass wir mit dem NVIDIA H100 HGX Server von Kaytus zum Erfolg des Projektes beitragen konnten. Das wissenschaftliche Ergebnis und die hohe Motivation des Teams ist beeindruckend.“

Hans-Peter Hellmann
CEO bei Delta Computer Products GmbH



Zugang zur modernsten Technologie am KI-Himmel.

Als Elite Partner von NVIDIA konnte Delta auf die Ressourcen des weltweit führenden Herstellers im Bereich KI zurückgreifen, was sich als äußerst vorteilhaft erwies. Die hochmoderne Technologie der H100 Server wurden optimal genutzt, um die Rechenkapazität des Neural Spacetime Labs zu maximieren. Die Server wurden bei Delta aufgebaut, getestet und anschließend betriebsbereit an die Hochschule Mannheim übergeben. Dort wurden sie mit Hochleistungs-Datenaufzeichnungsservern verbunden, um eine nahtlose Integration in die bestehende Infrastruktur sicherzustellen.

Die Partnerschaft mit Delta hat dem Institut für angewandte künstliche Intelligenz und Robotik eine erhebliche Ausweitung seiner Forschungsmöglichkeiten geboten. Die hochmodernen GPU-Server ermöglichen die Erzeugung umfangreicher Datensätze und das Training komplexer KI-Modelle, was zu wegweisenden Entwicklungen in der interventionellen Medizin und der Industrie führt.



„Die Firma Delta konnte tatsächlich die schwer verfügbaren H100 Server-Technologien bereitstellen und haben dabei noch die wirtschaftlich interessanteste Lösung angeboten.“

Prof. Dr. Marcus Vetter
Leiter des A²IR-Instituts

Das Ergebnis:

„Peter“ - Das Herzstück des A²IR

Peter, der Supercomputer im Neural Spacetime Lab, ist eine beeindruckende technologische Leistung. Er basiert auf der NVIDIA HGX™-KI-Supercomputing-Plattform von Kaytus und bietet eine KI-Leistung von 32 petaFLOPS FP8.

Peter ist mit 8 GPUs der NVIDIA H100 SXM-Baureihe ausgestattet, die jeweils 80GB GPU-Speicher haben. Dadurch kann er komplexe KI-Aufgaben effizient beschleunigen. Der integrierte NVIDIA NVLink Switch ermöglicht die Nutzung des gesamten 640GB GPU-Speichers für große KI-Modelle und Simulationen.

Zusätzlich verfügt Peter über 2 AMD EPYC CPUs mit jeweils 96 Kernen und 1,5 TB DDR5 RAM sowie einen internen NVMe SSD-Speicher von 15 TB für schnellen Zugriff auf lokale Daten. Zum Vergleich: Ein durchschnittlicher Laptop hat 16 GB RAM und etwa 500 GB SSD-Speicher.

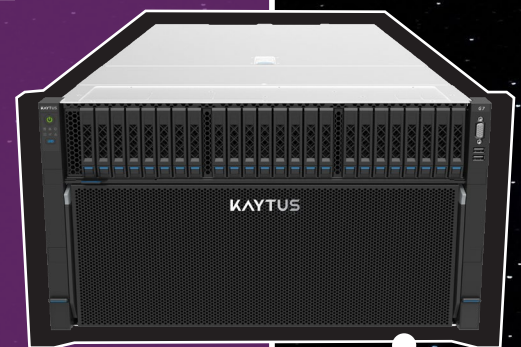
Dank des vorinstallierten Ubuntu Linux Betriebssystems kann Peter den vollständig optimierten NVIDIA KI- und HPC-Software-Stack aus dem NVIDIA NGC-Katalog nutzen. Diese Ausstattung macht Peter zu einem leistungsstarken Werkzeug für anspruchsvolle KI-Forschung und Entwicklung im Neural Spacetime Lab.

DELTA / Kaytus* KR 6288 - „Peter“

GPUs:	8x NVIDIA H100 SXM Tensor Core GPUs
GPU Memory:	(640 GB total (80 GB pro GPU))
CUDA® Cores (FP32):	135.168
CUDA® Cores (FP64):	67.584
Tensor Cores (TF32):	4.224
Performance:	32 petaFLOPS FP8
Interconnect GPUs:	NVLink Switch 900GB/s
CPU:	Dual AMD 9654 CPU, 96 Cores, 192 Threads, 192 Cores total, 2.4 GHz (Base) 3.7 GHz (Max Boost)
System Memory:	1.5 TB DDR5
Network:	2x 10G Ethernet (Dual Port)
Internal Storage:	15,3 TB (2x 7,68TB) U.3 NVME drives
Power Supplies/Usage:	4+4 redundant / 10,2 kW
Software:	Ubuntu Linux OS. with NVIDIA DL Frameworks

Es stehen weitere PCI und Disk Slots für Upgrades zur Verfügung. Diese Konfiguration ist optimiert für die vorgestellte Anwendung.

* Kaytus - ehemals bekannt als Inspur.



Eine neue Ära beginnt.

Neue KI-Assistenzsysteme in Sicht

Deep Learning basiert auf künstlichen neuronalen Netzen und der Verarbeitung großer Datenmengen. Es ermöglicht KI-Assistenzsystemen, Menschen bei Entscheidungen und Handlungen zu unterstützen. Diese Technologie erlaubt es Forschern, maschinelles Lernen und KI-Modelle zu entwickeln, die menschliche Interaktionen verstehen und darauf reagieren können.

Im Neural Spacetime Lab simulieren die Wissenschaftler des A²IR komplexe OP-Situationen. Sie nutzen dabei ein Netzwerk von 56 hochqualitativen Industriekameras mit Global-Shutter-Technologie, um diese Situationen aus verschiedenen Perspektiven zu erfassen. Mithilfe leistungsstarker GPUs und der NeRF-Technologie können die Forscher jede gewünschte Zeit und Position im Raum einnehmen – sogar für Positionen, die von keiner Kamera erfasst wurden. Dieser innovative Ansatz ist besonders nützlich für das Training von KI-Modellen im Bereich Computer Vision, erfordert jedoch eine enorme Rechenleistung.

„Die Qualität von KI-Assistenzsystemen hängt entscheidend von der Anzahl möglichst realistischer Daten ab. Mit dem Neural Space Time Lab haben wir nun die Möglichkeit, zeitlich hoch aufgelöst und räumlich hoch aufgelöst Realdaten zu erzeugen, mit denen wir anschließend die Modelle für KI-Assistenzsysteme trainieren können.“

Prof. Dr. Marcus Vetter

Die Zukunft der Medizin: Deep-Learning und Big-Data

Durch den Einsatz künstlicher Intelligenz und Big Data-Analyse können Mediziner bei medizinischen Eingriffen in ihrer täglichen Arbeit unterstützt werden. Diese Entwicklungen revolutionieren die Art und Weise, wie Mediziner Entscheidungen treffen, indem sie fallspezifische, statistische Informationen in ihre evidenzbasierten Vorgehensweisen einbeziehen.

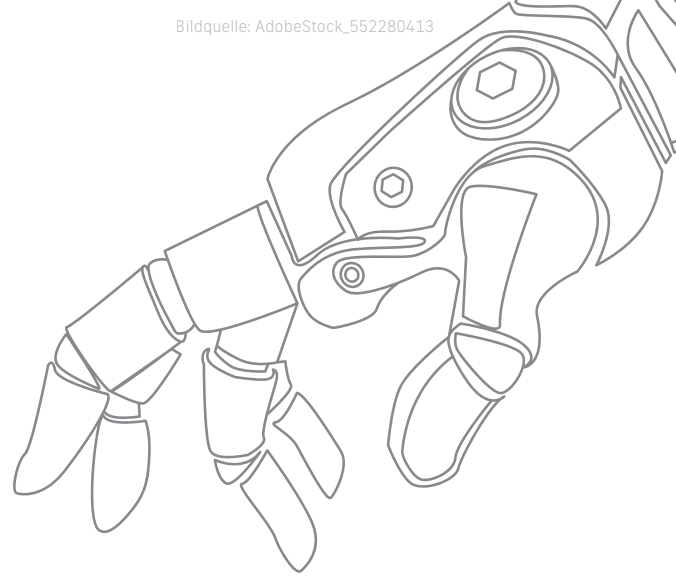


In den nächsten Jahren könnten KI-Assistenzsysteme während chirurgischer Eingriffe Empfehlungen zu alternativen Handlungsoptionen anbieten. Diese KI-gestützten Systeme können in Echtzeit Daten analysieren und dem Chirurgen wertvolle Hinweise zur Optimierung des Behandlungsverlaufs geben. Zudem ermöglicht die präzise Erkennung und Interpretation von Bildern, Texten und Mustern eine zukunftsorientierte Diagnose und Behandlungsplanung.

Durch die Analyse von Big Data kann die Technologie Abweichungen erkennen und frühzeitig Warnsignale für mögliche Komplikationen während medizinischer Eingriffe identifizieren. Darüber hinaus können komplexe Berechnungen dazu verwendet werden, die genaue Position von Personen, relevanten Gegenständen oder auch Organen zu bestimmen. Aus all diesen Informationen kann die KI ein holistisches Bild auswerten, was den Medizинern adäquat präsentiert werden kann.

Industrie 4.0:

Industrieroboter der nächsten Generation



Die Daten aus dem Neural Spacetime Lab am A²IR-Institut sind auch für die autonome Robotik wichtig. Zukünftig könnten autonome Transportsysteme in der Industrie menschliches Verhalten erkennen und darauf reagieren. Diese neuen Robotergenerationen bieten mehr Flexibilität und sind zuverlässiger.

Im Gegensatz zu herkömmlichen Systemen, die durch Infrarotsensoren eingeschränkt sind und bei unerwarteten Hindernissen Probleme haben, können diese fortschrittlichen Roboter weitgehend autonom handeln. Sie erkennen und umgehen Hindernisse selbstständig.

Dadurch können sie schnell und anpassungsfähig auf Veränderungen in ihrer Umgebung reagieren. Diese Vorteile sind besonders in dynamischen Umgebungen wie der Logistik bedeutsam. Das A²IR-Institut erforscht derzeit diese Anwendungen.

Diese Roboter sind nicht nur einfache Werkzeuge, sondern intelligente Partner in der Automatisierung der Industrie. Ihre Zuverlässigkeit und Anpassungsfähigkeit führen zu Zeit- und Kosteneinsparungen.



Mutig in die Zukunft. Synergien mit Weitblick

Diese fortschrittliche Servertechnologie bietet nicht nur Forschern, sondern auch Studierenden Zugang zu modernsten Technologien, um ihre Ausbildung zu optimieren. Die Studierenden der Hochschule Mannheim können im neuen Studiengang „KI-Ingenieurwissenschaften“ auf Hochleistungsservern eigene KI-Modelle trainieren.

Die erfolgreiche Kooperation zwischen Delta und dem A²IR zeigt eindrucksvoll, wie Technologieunternehmen und Forschungseinrichtungen gemeinsam künstliche Intelligenz verantwortungsvoll nutzen können. Deltas fortschrittliche Technologie und Expertise

haben es dem Labor ermöglicht, bedeutende Beiträge zur KI-Forschung zu leisten, die in der Industrie, Medizin und Logistik Anwendung finden und auch den Studierenden zugutekommen.

Der offene Umgang mit KI-Technologien und die Möglichkeit, eigene Ideen auszuprobieren, sind wesentliche Bestandteile der Philosophie des A²IR-Instituts in der Ausbildung der nächsten Generation von KI-Ingenieuren.

“Es macht einfach Spaß jetzt damit zu arbeiten und wir müssen schauen, wie wir den Zugang zum Server teilen, weil eigentlich jeder gleichzeitig darauf zugreifen möchte.”

Yannick Bukschat, Doktorand



Hier geht's zur Website!

**Noch mehr erfahren?
Kontaktieren Sie
uns für weitere
Informationen:**

 www.delta.de

 info@delta.de

 Tel.: +49 40 300 672 0



Bildquelle: AdobeStock_242280739

 **DELTA**
COMPUTER

Coming soon!

Neuer Stern in Sichtweite

NVIDIA B200 Tensor Core GPU



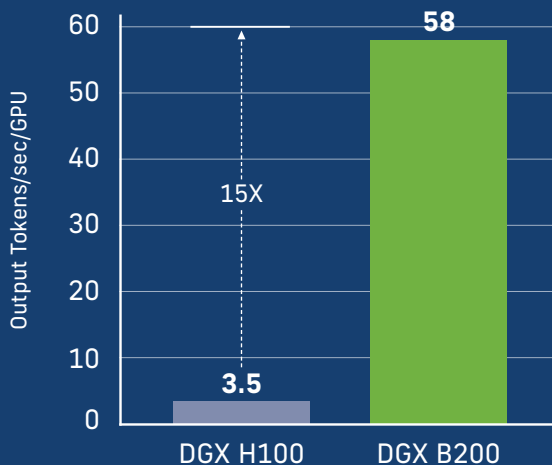
Superpower für die Forschung.
Delta solutions powered by NVIDIA.

NVIDIA Blackwell katapultiert Ihr Rechenzentrum in eine neue Ära des High-Speed-Computing und der generativen KI. Der angekündigte NVIDIA B200 Tensor Core Grafikprozessor basiert auf der Blackwell-Architektur und bietet einen enormen Fortschritt bei der Verbesserung von Inferenz-Workloads, wodurch Echtzeit-Performance für ressourcenintensive Sprachmodelle mit mehreren Billionen Parametern möglich wird.

Leistung neu definiert!

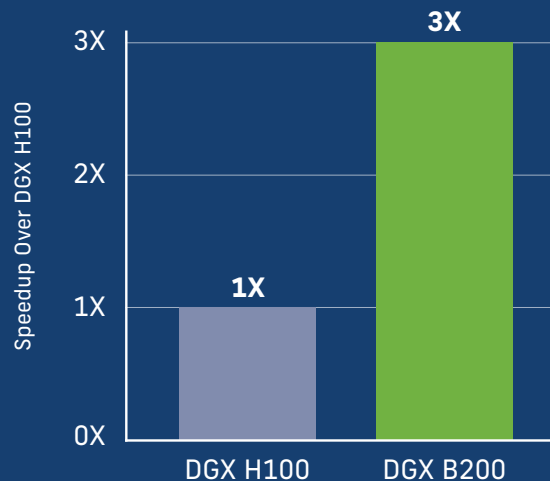
<https://www.deltacomputer.com/nvidia-dgx-b200>

Echtzeit-Inferenz für Large Language Models



GPT-MoE-1.8T Real-time Throughput

Neue Maßstäbe bei der KI-Trainingsleistung



GPT-MoE-1.8T Model Training Speed-Up