

9. Juli 2020

Schub für Big Data in der Wissenschaft

## Neuer Hochleistungsrechner des DLR in Jena eingeweiht

- Neuer Hochleistungsrechner in Jena stärkt die Kompetenzen des DLR im Bereich Machine Learning und Datenwissenschaften.
- 5856 Rechenkerne und ein 49 Terabyte großer Hauptspeicher verleihen dem HPDA-Cluster eine maximale Rechenleistung von rund 700 TFlops.
- Einweihung in Anwesenheit des Thüringer Wissenschaftsministers Wolfgang Tiefensee.
- Schwerpunkte: Raumfahrt, Digitalisierung, Klimawandel

Muster und Zusammenhänge in den immer größer werdenden Datenströmen unserer Zeit zu finden, ist eine große Chance und gleichzeitige Herausforderung der digitalen Gesellschaft. Wissenschaftlerinnen und Wissenschaftler des Instituts für Datenwissenschaften des Deutschen Zentrums für Luft- und Raumfahrt (DLR) in Jena erforschen wie aus der großen Menge an Messreihen aus dem All und auf der Erde bisher unbekannte Abhängigkeiten von Klimaparametern identifiziert, ungeahnte Beziehungen und Inhalte in Datenreihen nutzerorientiert visualisiert und Bürger im Rahmen von Citizen-Science-Projekten integriert werden können. Methoden des Machine Learning und Data Mining sind dabei die Werkzeuge der Forscher. Mit der am 9. Juli 2020 erfolgten Einweihung des High-Performance-Data-Analysis Clusters (HPDA) steht nun auch die passende Hochleistungsrechnerinfrastruktur bereit, um die Forschung zur Analyse von Big Data in der Wissenschaft einen weiteren Schub zu geben. Die Einweihung fand in Anwesenheit des Thüringer Ministers für Wirtschaft, Wissenschaft und Digitale Gesellschaft, Wolfgang Tiefensee, am Rechenzentrum der Friedrich-Schiller-Universität Jena statt.

„Dieser Hochleistungsrechner, der riesige Datenmengen und komplexe Operationen bewältigen kann, ist ein Kernstück der künftigen Forschungsarbeit am Institut“, sagte Thüringens Wissenschaftsminister Wolfgang Tiefensee. Eine solche Infrastruktur biete zugleich Ansatzpunkte für weitere intensive Kooperationen mit Forschungseinrichtungen der Region, insbesondere der Friedrich-Schiller-Universität Jena. Das Land habe die Investition in das HPDA-Cluster deshalb auch mit drei Millionen Euro unterstützt. „Drei Jahre nach Gründung ist das DLR-Institut auf einem guten Weg“, so Tiefensee weiter. Die heutige Inbetriebnahme sei ein weiterer wichtiger Meilenstein seiner Entwicklung. „Die Digitalisierung aller Lebensbereiche ist der Megatrend der Gegenwart. Auch dank des DLR-Instituts in Jena gestalten wir ein Stück dieser Zukunft von Thüringen aus mit.“

5856 Rechenkerne und ein 49 Terabyte großer Hauptspeicher verleihen dem HPDA-Cluster eine maximale Rechenleistung von rund 700 TFlops. "Mit diesem neuen Hochleistungsrechner stärken wir die Kompetenzen des DLR im Bereich Machine Learning und der Datenwissenschaften im Rahmen der Raumfahrtforschung", betonte DLR-Raumfahrtvorstand Prof. Hansjörg Dittus. "Das ist ein wichtiger Schritt zur Weiterentwicklung unseres Instituts für Datenwissenschaften in Jena, wo unsere

Wissenschaftlerinnen und Wissenschaftler nun technisch hervorragend unterstützt an leistungsfähigen Methoden zur Bearbeitung und Analyse großer Datenbanken forschen können“, so Dittus weiter. Wichtig ist auch die Zusammenarbeit der Jenaer Forscher mit weiteren DLR-Instituten. „Der HPDA-Cluster ist das zentrale Werkzeug für die Entwicklung neuer Methoden, die dann unter anderem von Instituten der Erdbeobachtung auf dem Cluster des Leibniz-Rechenzentrums im operationellen Betrieb genutzt werden“, ergänzt Dr. Robert Axmann, Gründungsdirektor des DLR-Instituts für Datenwissenschaften.

## **Data-Mining für die Klimaforschung**

In Jena weiterentwickelte Methoden des Data Mining und Machine Learning sollen zukünftig dazu beitragen in den Datenströmen der Klimaforschung bisher unerkannte oder unterschätzte Abhängigkeiten zwischen verschiedenen Klimaparametern zu identifizieren. Als eines der ersten Vorhaben wird das EU-Forschungsprojekt iMIRACLI (innovative Machine learning to constrain Aerosol-cloud CLimate Impacts) das neue HPDA-Cluster nutzen, um die Auswirkungen von Aerosol-Wolken-Wechselwirkungen von der mikroskopischen Skala bis zum großräumigen Klima zu verfolgen und zu quantifizieren. Aerosole entstehen unter anderem bei der Verbrennung fossiler Energieträger und sind fein verteilte, mikroskopisch kleine Partikel in der Luft, die die Wolkenbildung beeinflussen. Zudem streuen und absorbieren sie Sonnenstrahlung und beeinflussen so die Strahlungsbilanz der Atmosphäre und damit das Klima. Die Erde mit ihrer Atmosphäre, ihren Ozeanen und Kontinenten ist eines der am meisten beobachteten komplexen dynamischen Systeme. Satelliten und Bodenstationen liefern seit Jahrzehnten in immer kürzerer Zeit immer größere Datenmengen für mittlerweile fast alle Teile der Erde, die analysiert werden müssen und das ergänzt durch globale Klimamodelle, die die Dynamik der Atmosphäre und Ozeane simulieren und damit Projektionen für zukünftige Klimaentwicklungen liefern.

## **Visuelle Analyse und Citizen Science**

Zudem werden die Forscher in Jena den neuen HPDA-Cluster im Bereich Visuelle Analyse nutzen, um Data-Mining-Verfahren weiterzuentwickeln, die riesige Datenströme innovativ prozessieren, darin ungeahnte Beziehungen und Inhalte entdecken und diese nutzerorientiert visualisieren. Komplexe Zusammenhänge, statistische Unsicherheiten oder auch Besonderheiten in den Daten werden so für den Nutzer leicht sichtbar. Und auch die Forschung zur Integration von Citizen-Science-Daten in Analyseprozesse werden mittels des neuen Hochleistungsrechners vorangetrieben. Ein Beispiel sind hier Daten, die Bürger bezüglich der Bodenfeuchte oder über den Zustand der Vegetation in ihrem lokalen Umfeld erheben, die dann in die Kalibrierung satellitengestützter Erdbeobachtungsdaten einfließen. Die gleichberechtigte Zusammenarbeit von Bürgern und wissenschaftlichen Partnern im Rahmen von Citizen-Science-Projekten hat sich in den letzten Jahren zu einer gewinnbringenden Partnerschaft etabliert, der das DLR in Jena zusätzlich mit der Einrichtung eines bürgerwissenschaftlichen Labors ab Herbst 2020 Rechnung tragen wird.

## **Datenautobahn zwischen DLR und Universität Jena**

Der neu installierte HPDA-Cluster konnte mit Unterstützung der Friedrich-Schiller-Universität Jena im Rechenzentrum der Universität installiert werden. Der von der Firma MEGWARE

gelieferte und eingerichtete Hochleistungsrechner besteht aus 118 MEGWARE Rechenknoten und 4 GPU-Knoten mit jeweils 8 NVIDIA Tesla V100 GPUs. Diese sind mit skalierbaren Intel® Xeon® Platinum CPUs ausgestattet und dienen vor allem zur Unterstützung der wissenschaftlichen Arbeit im Bereich Klimainformatik und Machine Learning. Zwischen den Rechenknoten werden die Daten über ein Intel Omni-Path Netzwerk mit 100 Gigabit pro Sekunde (GBit/s) übertragen, die Anbindung an das Institut selbst wird über eine innerstädtische 10 GBit/s Glasfaser-Anbindung realisiert. Der HPDA-Cluster ist eine wichtige Grundlage für die gemeinsame Forschung verschiedener wissenschaftlicher Arbeitsgruppen des DLR und der thüringischen Universitäten. Das DLR-Institut für Datenwissenschaften in Jena beschäftigt derzeit rund 40 Mitarbeiter und 20 Studierende und wird sich zukünftig schrittweise erweitern.

## Kontakt

- Falk Dambowsky

### **Presseredaktion**

Deutsches Zentrum für Luft- und Raumfahrt (DLR)

### **Media | Relations**

- Dr. Robert Axmann

### **Gründungsdirektor**

Deutsches Zentrum für Luft- und Raumfahrt (DLR)

### **Institut für Datenwissenschaften**