

*Gemeinsame Presseinformation von Helmholtz Zentrum München, Max-Planck-Institut für Plasmaphysik, Deutsches Zentrum für Luft- und Raumfahrt, Technische Universität München und Ludwig-Maximilians-Universität München*

PI 10/18

5.11.2018

## **Ran an die Datenberge**

*Digitalisierte Forschung / neue Graduiertenschule für Data Science in München*

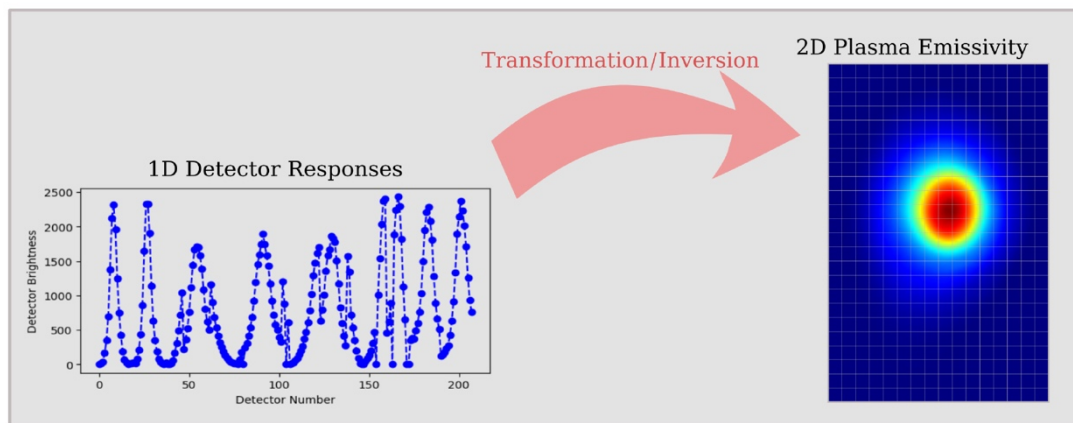
*Die digitalisierte Forschung liefert immer größere und komplexere Datenmengen. Sie bergen großes Potenzial etwa für die Biomedizin, die Energie- und Geoforschung oder Robotik, müssen aber auch bewältigt und interpretiert werden. Um die nächste Generation von Forschenden auszubilden, die diese sprichwörtlichen Big Data nutzbar machen, wurde die Munich School for Data Science @ Helmholtz, TUM & LMU (MuDS) ins Leben gerufen, die über sechs Jahre hinweg mit insgesamt zwölf Millionen Euro gefördert wird.*

Das Helmholtz Zentrum München, das Max-Planck-Institut für Plasmaphysik (IPP) und das Deutsche Zentrum für Luft- und Raumfahrt (DLR) sowie die Technische Universität München (TUM) und die Ludwig-Maximilians-Universität München (LMU) sind die Gründer der neuen Graduiertenschule. Darüber hinaus sind mit dem Leibniz-Rechenzentrum (LRZ) und der Max Planck Computing & Data Facility (MPCDF) zwei große Rechenzentren aus dem Raum München mit im Boot.

„Für große Herausforderungen braucht man große Lösungen: Wir sind froh, dass wir diese wichtigen Player im Großraum München für dieses Projekt zusammenbringen konnten“, erklärt Prof. Dr. Dr. Fabian Theis, Helmholtz Zentrum München/TUM, der das Projekt federführend koordinieren wird. Ziel des groß angelegten Vorhabens MuDS ist, die Ausbildung der nächsten Generation von Wissenschaftlerinnen und Wissenschaftlern im Bereich Data Science mit verschiedenen Domänenwissenschaften – Biomedizin, Plasmaphysik, Robotik und Erdbeobachtung – zusammen zu führen. Denn während die Datenvolumina und deren Komplexität stetig wachsen, fehlen Experten, um sie auszuwerten. Um es an einem Beispiel zu verdeutlichen: Jede einzelne Zelle in unserem Körper hat ein Erbgut von rund drei Milliarden Basenpaaren – das wäre eine Bibliothek von 3000 Büchern mit je 1000 Seiten, auf denen je 1000 Buchstaben stehen – und das ist nur eine Zelle.

Auch Beispiele aus den anderen Bereichen machen deutlich, wie sehr die Experten in Zukunft gefragt sind: Erdbeobachtungssatelliten der neuesten Generation liefern kontinuierlich Petabytes an Bild- und Messdaten für die Erforschung und das Management des globalen Wandels. Am IPP geht es unter anderem um die Modellierung künftiger Fusionsanlagen, für die modellbasierte Simulationen auf dem Computer ebenso nötig sind wie das Auswerten großer Datenmengen.

„Die neue Graduiertenschule“, sagt Prof. Dr. Frank Jenko, IPP-Sprecher in der MuDS, „wird dazu beitragen, diese Arbeiten auf ein neues Niveau zu heben, indem sie sowohl Plasmaphysiker und Datenspezialisten zusammenführt als auch den Austausch methodischer Anregungen unter den vier unterschiedlichen Anwendungsbereichen der MuDS ermöglicht.“



*Ein Beispiel für maschinelles Lernen in der Plasmaphysik: Das aus dem Plasma ausgesandte Röntgenlicht wird entlang von Sichtstrahlen von über 200 Detektoren beobachtet, die um das Plasma herum angeordnet sind. Die große aufgenommene Datenmenge (links) wird anschließend rechnerisch zu einem zweidimensionalen Bild des Plasmaquerschnitts zusammengesetzt (rechts). So wird erkennbar, wo genau in dem heißen Plasma die Röntgenstrahlung ausgesandt wurde. Für die schnelle rechnerische Transformation der Daten wurde ein speziell entwickeltes neuronales Netzwerk verwendet. (Abbildung: IPP, Florian Hendrich)*

Die Munich School of Data Science bietet bilaterale Projekte für Doktoranden an, die jeweils von zwei Partnern – einem domänenspezifischen Anwendungspartner und einem methodischen Partner – konzipiert werden. Dadurch wird sowohl eine methodische als auch anwendungsspezifische Ausbildung gewährleistet. Zudem wird ein auf die Teilnehmer zugeschnittenes Kursangebot mit einer detaillierten Onboarding-Phase und einer daran anschließenden vertiefenden Ausbildung angeboten. Das Trainingsprogramm wird in die bestehenden Kurse der Universitäten sowie der assoziierten Partner (LRZ und MPCDF) integriert, um eine zeitgemäße und exzellente Ausbildung zu gewährleisten. Angesiedelt ist die MuDS unter das Dach der bereits sehr erfolgreichen bestehenden Graduiertenschulen der Partner: HELENA, HEPP, TUM-GS und Munich Aerospace.

Das Fördervolumen beträgt über sechs Jahre hinweg insgesamt zwölf Millionen Euro und wird zur Hälfte von den beteiligten Institutionen bereitgestellt. Die andere Hälfte kommt von der Helmholtz-Gemeinschaft Deutscher Forschungszentren, in deren Strategie zur Digitalisierung der Forschung die School eingebettet ist (siehe Hintergrund). Die erste Ausschreibung für Doktoranden wird im Dezember 2018 unter <http://www.mu-ds.de> veröffentlicht.

### **Hintergrund:**

Die Helmholtz-Gemeinschaft investiert künftig pro Jahr 35 Millionen Euro in vier neue Plattformen zur Digitalisierung der Forschung. Sie sind jeweils an einem oder mehreren Helmholtz-Zentren verortet und schaffen ein aktives Netzwerk mit weiteren Forscherinnen und Forschern – dazu werden spezifische Förderlinien eingerichtet. Eine solche Plattform ist HIDA: Die Helmholtz-Plattform für wissenschaftlichen Nachwuchs und Graduiertenschulen im Bereich Information & Data Science. Innerhalb von HIDA werden fünf Graduiertenschulen an den Standorten Karlsruhe/Heidelberg, Jülich/Aachen/Köln, Hamburg, Berlin und eben München eingerichtet.