

Helmholtz “Information & Data Science Academy” (HIDA)

Autoren: Konrad (HZDR), Pfeiffenberger (AWI), Streit (KIT), Müller (GFZ), Dietrich (UFZ), Castell (HMGU)

In der heutigen Zeit ist der Umgang mit Informationen allgegenwärtig in der Wissenschaft und Wirtschaft und nimmt so einen elementaren Platz in unserem täglichen Handeln ein. Die digitale Transformation stellt eine entscheidende Herausforderung in allen wissenschaftlichen Disziplinen dar, sei es z.B. in der Gesundheitsforschung, Energieforschung oder Erdsystemforschung (aber ebenso in allen anderen Helmholtz Forschungsbereichen und Programmen), und hat einen starken sozialen und ökonomischen Einfluss auf die gesamte Menschheit. Digitalisierung ist daher ein kritischer Faktor, um Innovation am Standort Deutschland zu erhalten und weiter zu stärken. Im ihrem Kern steht die Transformation „von Daten zu Wissen“ durch eine synergetische Verbindung von Informatik, Mathematik, Sensortechnologie, Simulation und Daten-intensivem Rechnen mit Anwendungsfeldern aus dem breiten Spektrum der Natur- und Ingenieurwissenschaften, der Medizin und den Geistes- und Sozialwissenschaften. Und so werden zukünftig daten-intensive wissenschaftliche Methoden und Werkzeuge auch in allen Forschungsbereichen der Helmholtz Gemeinschaft Anwendung finden und zu neuen Erkenntnissen führen.

Dazu fehlt es jedoch an geeignetem Fachwissen und Fachpersonal in allen Helmholtz Zentren und Forschungsbereichen. Aufgrund der noch immer unzureichenden Fach- und Hochschulausbildung auf dem Gebiet Data Science sind Fachkräfte am Markt nicht in der nötigen Zahl zu bekommen. Da auch die Fachwissenschaftler das Handwerkszeug für den Umgang mit den Methoden und Werkzeugen der Data Science benötigen, ist eine Aus- und Weiterbildungsinitiative für die Helmholtz Gemeinschaft unbedingt notwendig.

Ziel der Helmholtz Information & Data Science Academy (HIDA) ist es, ein breites Wissen zu Methoden und Technologien von Data Science in der gesamten Helmholtz Gemeinschaft zu etablieren und gleichzeitig die internationale Sichtbarkeit und Strahlkraft der Helmholtz Gemeinschaft durch die HIDA zum Thema Data Science weiter zu steigern.

Weitere konkrete Ziele sind:

- Eine Community-Bildung für Data Science über die Helmholtz Gemeinschaft hinaus zu ermöglichen, ähnlich wie dies z.B. Oberwolfach für die Mathematik oder Schloss Dagstuhl für die Informatik erreichen,
- Synergien für eine wechselseitige Befruchtung nutzen und eine kritische Masse schaffen; diese Zielstellungen können vor allem aus der Verknüpfung der einzelnen Programme und Module erreicht werden,
- Impulse für die Vernetzung und Weiterentwicklung innerhalb und außerhalb der Helmholtz Gemeinschaft zu Data Science setzen und einen Beitrag leisten, um eine praxisnahe Lehre für das Thema Data Science an den Universitäten und Fachhochschulen zu entwickeln; regionale Research Schools können dabei als „Academy Sites“ wirken,
- Raum für Veranstaltungen rund um Data Science und ein Forum für Nutzer und Experten zur Erarbeitung konkreter Lösungsansätze zu schaffen,
- Die HIDA als Marke zu etablieren, die eine Referenz auf dem Gebiet Information und Data Science weit über die Helmholtz Gemeinschaft und Deutschland darstellt.

Gruppirt nach den **Hauptaufgaben** der Akademie können schrittweise folgende Instrumente entwickelt werden:

A. Think Tank und "Ideenschmiede"

- o Strategie- und Kreativ-Workshops
- o Beratung/Diskussion zu Berufsbildern und Curricula
- o Politische und politik-beratende Workshops

B. Vernetzungs- und Koordinierungsstelle

- o Trainee-Netzwerk (siehe Konzept HOCUS - Helmholtz Computational Science Exchange Program):
 - Austausch von Expertise zwischen Gruppen verschiedener Zentren
 - Experten (Fellow) einladen oder Besuch einer Gruppe zum Lernen (Trainee)
 - verfügbar für kurz- und mittelfristige Besuche (max. 3 Monate)
 - verfügbar für alle Karrierestufen
 - Trainee Network deckt die Reisekosten für die Besucher ab
- o Vernetzung mit regionalen Research Schools (im Sinne von Graduiertenschulen), die an den HGF-Zentren als „Academy Sites“ wirken:
 - Aufbau dieser Schulen mit regionalen Partnern (Universitäten)
 - die operative Steuerung der Schools und die Auswahl der Doktoranden bleibt dabei in der Hand der lokalen Partner,
 - Spezialisierung der Schulen entsprechend der lokalen Kompetenzen (u.a. durch „nationale Innovation Hubs“),
 - Gemeinsame Tagungen zum Interagieren bzgl. fachlichen und überfachlichen Themen --> siehe Begegnungsstätte
 - Young Data Scientist Award

C. Aus- und Weiterbildung

- o Angebote von fachspezifischen Kursen, Semesterangeboten und Vorträgen weltweit führender Experten
- o Ort/Rahmen für Sommerschulen und Graduiertenschulen (jedoch keine separate Grad-School)
- o Teile eines dualen Studiums (IHK-Ausbildung mit „FH-Master“)
- o Moderne Formate, z.B. Hackathons, MOOCs, TEDex@HGF
- o Themenschwerpunkte:
 - Daten-Analyse und -Auswertung
 - Maschine Learning / Deep Learning
 - Modellierung und Optimierung
 - Visualisierung und Bildverarbeitung
 - Verschlüsselung, Sicherheit und Vertrauenswürdigkeit
 - Softwareentwicklung und Qualitätssicherung
 - Management und Verfügbarkeit von Daten (nach FAIR-Prinzipien)
 - Aufbau und Nutzung von Infrastrukturen

D. Begegnungsstätte

- o Thematische Workshops
- o Ort/Rahmen für Sommerschulen und Seminare von Graduiertenschulen (z.B. im Rahmen der Academy Sites)
- o Platz für Sabbatical (längere Treffen von Experten)
- o Raum zum gemeinsamen „Experimentieren / Ausprobieren von Ideen“

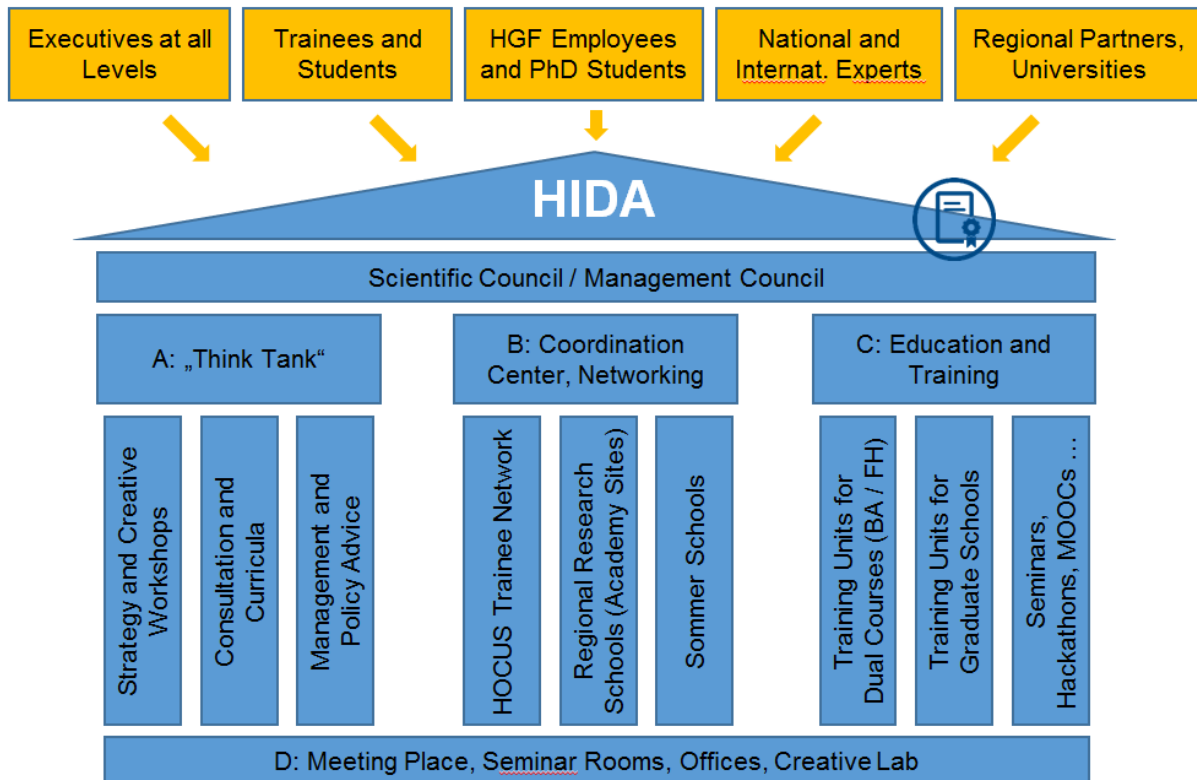


Bild 1: HIDA Übersicht

Zielgruppen der HSDA umfassen u.a.:

- Helmholtz Mitarbeiterinnen und Mitarbeiter (Wissenschaftler, technisches und administratives Personal) und PhD Studenten
- Auszubildende und Studenten (Schwerpunkt BA / FH)
- Experten aus dem nationalen und internationalen Umfeld
- Führungskräfte auf allen Ebenen der Helmholtz Gemeinschaft
- Regionale Partner, Universitäten

Rahmenbedingungen

- Unabhängiger wissenschaftlicher Beirat als „Programmkomitee“
- Offene Verfahren zum Einbringen von Themen
- Genügend Raum:
 - o Seminarräume für ca. 30-50 Personen
 - o ca. 3-5 Büros für temporäre Nutzung
 - o Übernachtungsmöglichkeiten
 - o sehr gute Erreichbarkeit
- Orte und Möglichkeiten für agiles Arbeiten & Leben
- Stipendien zur Förderung von Teilnahmen (von Teilnehmern wie auch Experten)
- Familienfreundliche Atmosphäre insb. für längere Aufenthalte

Aufgrund der Komplexität der Aufgabe wird ein mehrstufiges Vorgehen vorgeschlagen. In Phase 1 werden vor allem die Aufgabengebiete Think Tank (A) und Vernetzung (B) adressiert. In Phase 2 soll die Aus- und Weiterbildung starten und geeignete Räumlichkeiten gefunden werden und in Phase 3 wird die HIDA langfristig konsolidiert.

Konkrete nächste Schritte dafür sind:

- **Phase 1: Conceptual Design und Vernetzungsaktivitäten**
 - o Kernteam etablieren, um das konzeptionelle Design bzw. den Förderantrag einschließlich detaillierter Kosten zu entwickeln, dieses könnte z.B. nach dem Modell des Open Science Büros zentral an einem Zentrum oder der GS angesiedelt werden
 - o Organisation und Koordinierung der Aktivitäten zu Trainee Networks und zur Etablierung regionaler Research Schools bereits in der Aufbauphase
 - o Aufwand/Kosten/Ressourcen zur Erarbeitung eines CDR:
 - Laufzeit: 1 Jahr
 - Personal: 2 Personen (z.B. eine Person mit eher kaufmännischem Hintergrund und die andere Person mit wissenschaftlichem Hintergrund) + Back-office-Unterstützung (ca. 150 T€/a)
 - Budget für Trainee Netzwerk (z.B. 250T€/a)
- **Phase 2: Implementierung und Etablierung**
 - o Kernteam verantwortet die Organisation, PR und Event Planung
 - o Vollkonzept, ggf. auch mit „fester Adresse“ (eigenem Haus) unter Einbeziehung aller Zentren
 - o Etablierung des wissenschaftlichen Beirates als „Programmkomitee“ (könnte im ersten Schritt der Inkubator „Information & Data Science“ sein)
 - o Definition von Expertenteams, Inhalten und „Curricula“ (aber nicht im Sinne von Studiengängen gemeint)
 - o Suchen einer geeigneten Lokation
 - o Laufzeit: 2-3 Jahre
- **Phase 3: Konsolidierung**
 - o Fortführen bestehender Aktivitäten
 - o Anpassung der Inhalte an Bedarfe
 - o Einbeziehung weiterer Akteure