

HELMHOLTZ-INKUBATOR INFORMATION & DATA SCIENCE

Weiterentwicklung eines neuartigen, gemeinschaftsweiten Ansatzes

INHALT

1. Einleitung	3
2. Information & Data Science in der Helmholtz-Gemeinschaft	3
3. Der Helmholtz-Inkubator Information & Data Science	5
4. Der Helmholtz-Inkubator als zukunftsweisender Katalysator	6
5. Vision Information & Data Science	9

1. EINLEITUNG

Die **digitale Transformation** ist die größte Herausforderung für Wissenschaft, Wirtschaft und Gesellschaft zu Beginn des 21. Jahrhunderts. Sie bietet **enorme Chancen** in nahezu allen Bereichen des Lebens – innovative Formen von Arbeit und Zusammenleben, völlig neuartige Plattformen für Handel und Wissenschaft, ungeahnte Möglichkeiten für die Medizin, revolutionäre Ansätze für Netzwerke der Wasser- und Energieversorgung, zukunftsweisende Verfahren für Klimaschutz und Migration und vieles mehr.

Im Kern der digitalen Transformation steht die Wertschöpfungs- und Erkenntniskette „**von Daten zu Wissen zu Innovation**“. Dies ist auch die entscheidende Herausforderung in allen wissenschaftlichen Disziplinen, welche die Grundlagen für den gesellschaftlichen Fortschritt der Zukunft bereiten.

Damit Deutschland diese Entwicklung mitgestalten und daran partizipieren kann, sind ambitionierte, mutige und zukunftsweisende Schritte erforderlich. Die Verbindung von Informatik, Mathematik, Statistik, Sensortechnologie, Simulation und datenintensivem Rechnen mit anspruchsvollen Anwendungsfeldern aus dem breiten Spektrum der Natur- und Ingenieurwissenschaften, der Medizin sowie den Geistes- und Sozialwissenschaften wird ein dynamischer Motor des Innovations- und Forschungsstandortes Deutschland sein. Diese innovativen Verbindungen werden völlig neue wissenschaftliche Erkenntnisse mit erheblichen Mehrwerten für Wirtschaft und Gesellschaft hervorbringen.

Deutschland hat im Feld der Hochtechnologien mit großem transformativem Potenzial eine sehr gute Ausgangsposition im weltweiten Vergleich. Um diese Position zu halten und auszubauen, sind neben gesamtgesellschaftlichen Anstrengungen auch **völlig neuartige Ansätze im Wissenschaftssystem** erforderlich.

2. INFORMATION & DATA SCIENCE IN DER HELMHOLTZ-GEMEINSCHAFT

Die **Helmholtz-Gemeinschaft Deutscher Forschungszentren** leistet Beiträge zur Lösung großer und drängender Fragen von Gesellschaft, Wissenschaft und Wirtschaft durch wissenschaftliche Spitzenleistungen in sechs Forschungsbereichen: Energie, Erde und Umwelt, Gesundheit, Schlüsseltechnologien/ Information, Materie sowie Luftfahrt, Raumfahrt und Verkehr. Sie ist mit über 38.000 Mitarbeiterinnen und Mitarbeitern in 18 Forschungszentren die größte Wissenschaftsorganisation Deutschlands.

Als Forschungsorganisation, die sich zum Ziel gesetzt hat große und aktuelle gesellschaftliche Herausforderungen zu adressieren, hat die Helmholtz-Gemeinschaft in den letzten Jahrzehnten ein enormes Kompetenzportfolio im Bereich **Information & Data Science** aufgebaut: auf Gebieten wie Informationsverarbeitung, Big Data, Data Analytics, Simulation, Modellierung, Bioinformatik, bildgebenden Verfahren, Forschungsdaten-Management, High Performance Computing, Robotik, technischen sowie biologischen Informationssystemen und vielen weiteren Technologien. Wie kaum eine andere Forschungsorganisation verfügt sie über einen exponentiell wachsenden Schatz von Big Data.

Die Helmholtz-Zentren und Forschungsbereiche haben auch eine herausragende Ausgangsposition für eine erfolgreiche, synergetische Verbindung der Kompetenzen: als Betreiber großer Forschungsinfrastrukturen (beispielsweise Satellitenmissionen oder Großanlagen der Kern- und Teilchenphysik), als Anwender von Supercomputing der neuesten Generation sowie als Kompetenzträger für komplexe Simulationen (beispielsweise umfangreicher Erd- und Klimamodelle, virtuelles Material-Design und Systembiologie). Sie sind daher auch schon seit Langem mit dem gesamten Data-Lifecycle vertraut: der Forschungsplanung, dem Erheben, der Handhabung und Pflege, dem Analysieren, dem Auswerten und der Nutzbarmachung sehr großer und komplexer Datenmengen. An allen Standorten gibt es teilweise weltweit führende Ansätze und herausragende Methodenkompetenz in den jeweiligen Teilgebieten.

Darauf aufbauend und darüber hinaus gehend verstärkt die Helmholtz-Gemeinschaft zurzeit und zukünftig im hochaktuellen Themenfeld Information & Data Science die eigene Kompetenz, schafft Synergien in der Forschungslandschaft, greift die Entwicklungen im nationalen, europäischen und internationalen Kontext auf und setzt neue disruptive Ansätze um. Neben der Stärkung der Einzeldisziplinen mit modernsten daten- und informationswissenschaftlichen Methoden ist es erklärtes Ziel, das Thema Information & Data Science auf Gemeinschaftsebene disziplinübergreifend voranzutreiben.

**Die Helmholtz-Gemeinschaft adressiert das komplexe Themenfeld
Information & Data Science auf allen Ebenen:**

- Es hat hohe Priorität in der Agenda des Präsidenten.
- Zentren und Forschungsbereiche berücksichtigen in ihren jeweiligen Strategien das große Potenzial dieses Themenfeldes.
- Im Rahmen der Neuausrichtung der Forschungsbereiche wird der bisherige Forschungsbereich Schlüsseltechnologien in einen neuen Forschungsbereich Information weiterentwickelt.
- Die Gemeinschaft errichtete mehrere neue Institute zu Simulations- und Datentechnologien und Cybersicherheit und plant aktuell die Aufnahme eines neuen Helmholtz-Zentrums für Informationssicherheit am Standort Saarbrücken.
- Wissenschaftlicher Nachwuchs wird in diesem Bereich auf neuen Wegen und in großer Zahl in einem neuen Netzwerk regionaler Helmholtz Information & Data Science Schools ausgebildet.
- Der vom Präsidenten der Helmholtz-Gemeinschaft initiierte Helmholtz-Inkubator „Information & Data Science“ potenziert die einzelnen Initiativen als ein neuartiger, gemeinschaftsweiter Think-Tank und Zukunftsmotor.

3. DER HELMHOLTZ-INKUBATOR INFORMATION & DATA SCIENCE

Der **Helmholtz-Inkubator Information & Data Science** wurde im Juni 2016 von der Gemeinschaft ins Leben gerufen, um die vielfältige, dezentrale Expertise der Helmholtz-Gemeinschaft intelligent zusammenzuführen. Dazu haben alle Helmholtz-Zentren je zwei hochkarätige Wissenschaftlerinnen und Wissenschaftler in den Helmholtz-Inkubator entsandt. Diese 36 Fachleute vertreten zusammen eine enorme fachliche Breite und decken viele innovative Kompetenzen ab. Unterstützt werden sie durch ausgewiesene Experten aus namhaften, forschenden Unternehmen.

Der Helmholtz-Inkubator integriert bestehende, zukunftsweisende Initiativen der Helmholtz-Gemeinschaft durch einen **gemeinschaftsweiten Bottom-up-Prozess**. Die regelmäßige Zusammenführung und Verdichtung der Expertise unserer Zentren ermöglicht die visionäre Gestaltung des Themas Information & Data Science. Dabei geht der Helmholtz-Inkubator völlig neue Wege, um das Zukunftsfeld kraftvoll, umfangreich und fachübergreifend zu gestalten.

Der Helmholtz-Inkubator verfolgt derzeit **folgende Ziele**:

- die regelmäßige Interaktion kreativer Köpfe aus der gesamten Gemeinschaft,
- die Schaffung von Grundlagen für innovative, interdisziplinäre Netzwerke und Ansätze,
- die Identifizierung zukunftsweisender Themenfelder und disruptiver Pilotprojekte,
- die Planung von langfristig angelegten, systemischen und strukturellen Ergänzungen.

Für die Unterstützung innovativer Aktivitäten und vielversprechender Pilotprojekte hat der Präsident 30 Mio. Euro aus dem Impuls- und Vernetzungsfonds der Helmholtz-Gemeinschaft bereitgestellt; mit der Ko-Finanzierung der Helmholtz-Zentren werden dadurch umgehend **mehr als 60 Mio. Euro für das Zukunftsthema mobilisiert**. Die Zuweisung erfolgt in einem strengen Peer-Review-Verfahren mit renommierten, internationalen Expertinnen und Experten.

4. DER HELMHOLTZ-INKUBATOR ALS ZUKUNFTSWEISENDER KATALYSATOR

Mit der konsequenten Umsetzung dieser Ziele setzt der Helmholtz-Inkubator **gemeinschaftsweite Impulse** im Bereich Information & Data Science, die erstens kontinuierlich zur Bearbeitung neuer Themen beitragen, zweitens zu essentiellen strukturellen Initiativen weiterentwickelt werden, und drittens Grundlagen zur langfristigen Bearbeitung groß-skaliger Forschungsthemen schaffen. Diese Aktivitäten können Anknüpfungspunkte sein, um Interaktionen und Strukturen weit über die Helmholtz-Gemeinschaft hinaus aufzubauen und das enorme Potenzial gemeinsam zu verwirklichen.

PILOTPROJEKTE

Der Helmholtz-Inkubator trat im Oktober und Dezember 2016 sowie im Mai 2017 zu mehrtägigen Beratungen zusammen. Die Expertinnen und Experten haben im Dezember 2016 **neuartige, fach- und disziplinübergreifende Forschungsschwerpunkte** identifiziert. Darauf aufbauend erarbeiteten Forscher-Konsortien Anträge für disruptive Pilotprojekte, die einen zukunftsweisenden Impuls im Bereich Information & Data Science geben und neue interdisziplinäre Zusammenarbeit ermöglichen. **Die ersten fünf Pilotprojekte** wurden im Juni 2017 ausgewählt und haben ein Fördervolumen von 18 Mio. Euro. Die Pilotprojekte sind bereits in der Umsetzung, entwickeln neueste Analysetechniken und beschreiten völlig neue Wege, um aus Daten Wissen zu generieren. Weitere zukunftsweisende Pilotprojekte werden zügig folgen.

Ausgewählte Pilotprojekte

*Im „**Helmholtz Analytics Framework**“ werden hochmoderne, grundlegende Analysetechniken in einem Co-Design-Ansatz mit Domänenwissenschaftlern und Informationsexperten entwickelt. Dies ermöglicht auch eine frühzeitige Ausbreitung der Methoden.*

*Das Pilotprojekt „**Sparse2Big**“ stellt methodische Grundlagen für den Umgang mit noch spärlichen, teilweise fehlerhaften, großen Datensätzen bereit: Durch Vervollständigung und robuste Modellierung wird Big Data hoher Qualität, insb. für die Gesundheitsforschung, geschaffen.*

*In den „**Reduced Complexity Models**“ werden intelligent reduzierte Modelle von komplexen Computersimulationen entwickelt. Diese ermöglichen den Umgang mit enormen Datenmengen aus komplexen, mathematischen Computermodellen.*

*Im Rahmen des Projekts „**Automated Scientific Discovery**“ werden innovative und zuverlässige Methoden der künstlichen Intelligenz genutzt, um Beziehungen in großen Mengen wissenschaftlicher Daten automatisiert zu erforschen.*

*Das Pilotprojekt „**Imaging at the Limit**“ entwickelt interdisziplinäre Verfahren der Bildrekonstruktion, die Messungen in Bilder überführen. Dies erhöht die allgemeine Informationsgewinnung, denn Bildgebungsverfahren stellen eine essentielle Informationsquelle in nahezu jedem Forschungsgebiet dar.*

UMFASSENDE ANSÄTZE FÜR STRUKTURELLE HERAUSFORDERUNGEN

Im Verlauf des ersten Inkubator-Workshops im Oktober 2016 identifizierten die ExpertInnen **zentrale, strukturelle Herausforderungen** für die Helmholtz-Gemeinschaft und die gesamte Forschungslandschaft. Diese wurden im Frühjahr 2017 in mehreren fokussierten Arbeitsgruppen konkretisiert und zu schlagkräftigen Zukunftskonzepten entwickelt. Diese Ergebnisse der Arbeitsgruppen wurden in einem Inkubator-Workshop mit allen Expertinnen und Experten im Mai 2017 diskutiert und geschärft.

Die wissenschaftlichen und kaufmännischen Vorstände der Helmholtz-Gemeinschaft haben sich in einer **außerordentlichen, fachlichen Sitzung aller Vorstände** im September 2017 eingehend mit diesen fünf strukturellen Herausforderungen befasst. Sie haben diese zur detaillierten Ausarbeitung empfohlen, um eine potenzielle gemeinschaftsweite, nachhaltige Umsetzung vorzubereiten. Die Arbeitsgruppen des Helmholtz-Inkubators haben im Oktober und November 2017 die weitere Konkretisierung vorgenommen und erstellen bis Mitte 2018 **umfassende Konzepte für nachhaltige, strukturelle Ergänzungen** für die Helmholtz-Gemeinschaft, gegliedert in die folgenden fünf Handlungsfelder.

Wissenschaftlicher Nachwuchs

Information & Data Science ist einer der Schlüssel für Forschung und Entwicklung an der Weltspitze. In Relation zur Größe der sich auftuenden Chancen und Herausforderungen ist jedoch ein Mangel an geeignetem Fachpersonal in Wissenschaft, Forschung und Entwicklung zu konstatieren. Insbesondere fehlt es an Talenten, die über profunde Kenntnisse sowohl in modernen Informationstechnologien als auch in einer dezidierten wissenschaftlichen Domäne verfügen. Ziel ist es daher, massiv in ein breit aufgestelltes Netzwerk postgradualer Ausbildung in diesem Bereich zu investieren und ein neuartiges Ökosystem mit kritischer Masse an Forschern der Information & Data Science sowie Domänenforschern zu entwickeln. Dieses Netzwerk wird aus einem halben Dutzend dedizierter,

regionaler Helmholtz Information & Data Science Schools bestehen, die derzeit in enger Zusammenarbeit mit Partner-Universitäten ausgestaltet werden, und als Helmholtz Information & Data Science Academy (HIDA) die wichtigsten einschlägigen Ausbildungsaktivitäten mehrwertstiftend verbinden.

Maschinelles Lernen

Die Entwicklung von Künstlicher Intelligenz – und spezifischen, zukunftsweisenden Verfahren des Machine Learnings sowie Deep Learnings – wirken bereits disruptiv in der gesamten Forschungslandschaft. Die Verfahren ermöglichen die Beantwortung völlig neuer wissenschaftlicher Fragestellungen und eine beachtliche Verbesserung von Forschung in quantitativer sowie qualitativer Weise. Um das enorme Potenzial für die eigene Forschung und die Gesellschaft noch weiter nutzbar zu machen, entwickelt die Helmholtz-Gemeinschaft wegweisende Technologie-Plattformen, auf denen sie ihre enorme Expertise in der Methodenentwicklung und in den Domänenwissenschaften bündelt und die eigenen Expertinnen und Experten mit weiteren Partnern aus dem Wissenschaftssystem zusammenführt.

Bildgebende Verfahren

Bilderkennung und bildgebende Verfahren sind von essentieller Bedeutung für sämtliche Felder moderner Forschung und Entwicklung, da ein immer größerer Teil der Informationen in Bildform erhoben und verarbeitet wird. Die Helmholtz-Gemeinschaft plant daher die systematische Weiterentwicklung und konsequente Nutzbarmachung neuester Ansätze, um völlig neue Erkenntnisse in vielen Forschungsfeldern zu ermöglichen.

Mehrwerte aus Forschungsdaten durch Metadaten

Ein leistungsstarkes und zukunftsicheres Forschungsdatenmanagement ist von essentieller Bedeutung, denn Forscher erheben immer vielfältigere und höchst komplexe Datensätze. Um aus den Daten Wissen zu generieren, die Ergebnisse zu reproduzieren oder sie nachnutzen zu können (FAIR-Prinzipien), müssen der Entstehungsprozess und der Kontext bestmöglich dokumentiert sein. Basierend auf ersten Initiativen in diesem Bereich verstärkt die Helmholtz-Gemeinschaft daher ihre Kompetenzen in der Erhebung, dem Management und der Analyse von Metadaten. Damit kann die Gemeinschaft auch einen prominenten Beitrag zum Aufbau von effektiven nationalen, europäischen und internationalen Forschungsdateninfrastrukturen (NDFI, EOSC u.w.) leisten.

Basistechnologien und grundlegende Dienste

Die Erforschung großer gesellschaftlicher Herausforderungen erfolgt zunehmend in dezentralen, dynamischen Forscher-Netzwerken. Weiterhin werden enorme Datenmengen an unterschiedlichen Standorten erhoben, ausgewertet und gespeichert. Um moderne Forschungsk Kooperationen bestmöglich zu unterstützen, ist ein leistungsfähiges Forschungsnetzwerk mit unterstützenden Dienstleistungen (z.B. clouds, remote computing, authentication networks etc.) erforderlich. Die Helmholtz-Gemeinschaft hat hier bereits über Jahrzehnte gewachsene Kompetenzen und Kapazitäten und wird die Entwicklung dieser „core technologies“ systematisch für die gesamte Gemeinschaft und ihre Partner bündeln. Dabei werden wir auch auf die Vorleistungen der bereits initiierten Helmholtz Data Federation (HDF) zurückgreifen und auch diese passgenau weiterentwickeln.

5. VISION INFORMATION & DATA SCIENCE

NÄCHSTE SCHRITTE

Die ersten konkreten Ergebnisse und Projekte des Helmholtz-Inkubators belegen das **enorme Potenzial** dieses Ansatzes. Die Pilotprojekte setzen zukunftsweisende Impulse und realisieren völlig neuartige Formen der interdisziplinären Zusammenarbeit. Der Helmholtz-Inkubator ermöglicht es uns, die fundamentalen Herausforderungen im Bereich Information & Data Science umfassend, wissenschaftsgeleitet und gemeinsam mit Partnern anzugehen. Interne und externe Experten analysieren die wissenschaftlichen Bedarfe und entwickeln in diesem ambitionierten Prozess langfristig tragfähige Handlungsoptionen. Das Top-Management der Helmholtz-Gemeinschaft richtet diese gemeinschaftsweiten Handlungsoptionen strategisch aus. Eine kontinuierliche Mitwirkung aller Ebenen ermöglicht eine bedarfsorientierte, konzertierte Umsetzung nach detaillierter und umfassender Planung. So setzt die Helmholtz-Gemeinschaft als gestaltender Akteur kontinuierlich **neue Impulse** und entwickelt nachhaltige, **tragfähige Ergänzungen** für die Forschungslandschaft.

In der näheren Zukunft sind die folgenden **konkreten Schritte** geplant:

- Bis Mitte 2018 werden fünf Konzepte für die oben umrissenen, strukturellen Ergänzungen ausgearbeitet. Daran sind der gesamte Helmholtz-Inkubator, seine Arbeitsgruppen mit weiteren Wissenschaftlerinnen und Wissenschaftlern sowie externe Expertinnen und Experten in einem dynamisch-konzertierten Prozess beteiligt. Im September 2018 sollen diese Konzepte den Vorständen der Gemeinschaft unterbreitet und zur Umsetzung empfohlen werden.
- Die von der gesamten Gemeinschaft getragenen Konzepte sollen im Anschluss zügig realisiert und nach Abschluss einer ersten Aufbauphase sämtlichen Partnern im deutschen Wissenschaftssystem als Angebot zur Kooperation und Teilhabe vorgelegt werden.
- Der Helmholtz-Inkubator wird fortlaufend die führenden Expertinnen und Experten der Helmholtz-Gemeinschaft versammeln, um neue Ansätze für disruptive Forschungsprojekte zu entwickeln. Diese neuen Pilotprojekte werden die bereits etablierten Vorhaben ergänzen und Information & Data Science-Forschung kontinuierlich mit neuen Herangehensweisen erweitern.
- Weitere Inkubator-Workshops sollen die konsequente Vernetzung der Aktivitäten im Themenfeld Information & Data Science voran bringen, das neuartige Ökosystem weiter strukturell ausbauen sowie immer neue Angebote an Wissenschaft und Gesellschaft entwickeln.

VISION

Die Helmholtz-Gemeinschaft verfolgt gemeinschaftsweite, visionäre Anstrengungen, um das Zukunftsthema Information & Data Science in der erforderlichen Ganzheit voranzutreiben. Der Helmholtz-Inkubator agiert als **Think-Tank**, der die Helmholtz-Gemeinschaft als Innovationsmotor auch in Zukunft kraftvoll unterstützt. Um die vielfältigen Anforderungen zu bewältigen, soll er kontinuierlich und langfristig weiterentwickelt werden, aktuelle Trends aufgreifen und neue Perspektiven berücksichtigen.

Der Helmholtz-Inkubator baut an einer **innergemeinschaftlichen Community mit kritischer Masse** an Forschern der Information & Data Science sowie an Domänenforschern aus, um die kontinuierliche Rückspiegelung neuer Ansätze in die Einzeldisziplinen zu gewährleisten und den Methoden- und Wissenstransfer verstärkt über Köpfe zu fördern. Dieser neuartige Ansatz sichert die Spitzenposition der Forschung im nationalen und internationalen Wettbewerb.

Die Aktivitäten des Helmholtz-Inkubators bieten mannigfache Anknüpfungspunkte für verschiedene Partner und können durch weitere Kompetenzträger symbiotisch ergänzt werden. Sie bilden einen tragfähigen und kraftvollen Nukleus für **groß-skalige Forschungsvorhaben** im nationalen sowie europäischen Kontext. Der Helmholtz-Inkubator soll zu einer Plattform werden, um Interaktionen weit über die Helmholtz-Gemeinschaft hinaus aufzubauen und das enorme Potenzial der Zukunftsthemen gemeinsam zu verwirklichen. Die Helmholtz-Gemeinschaft und ihre Partner sind entschlossen, auf diesem Weg auch wichtige Beiträge zur Gestaltung einer Nationalen Forschungsdateninfrastruktur zu leisten.

Zukunftsweisende Forschung erfordert die **internationale Kooperation** mit renommierten Expertinnen und Experten aus aller Welt. Gleiches gilt für die Gewinnung und Bindung exzellenter Forscherinnen und Forscher sowie exzellenten Nachwuchses. Der Helmholtz-Inkubator unterstützt durch seine Programme die weltweite Vernetzung des deutschen Forschungsstandorts mit attraktiven Partnern im Feld Information & Data Science nachhaltig.

Mit dem Helmholtz-Inkubator setzt die Helmholtz-Gemeinschaft – basierend auf ihren vielfältigen Stärken – völlig neue strategische Impulse zur **Erschließung zukunftsweisender, interdisziplinärer Forschungsfelder**. Er wird schlagkräftige Konzepte und Handlungsoptionen entwickeln, um belastbare Vorschläge zur Allokation weiterer Fördermittel zu unterbreiten.

Schließlich positioniert der Helmholtz-Inkubator die **Helmholtz-Gemeinschaft als gestaltenden Akteur** auf dem Feld Information & Data Science. Er setzt entscheidende Impulse für die Aufstellung der Helmholtz-Gemeinschaft in diesem Zukunftsthema und strahlt damit auf die deutsche und internationale Forschungslandschaft aus.

