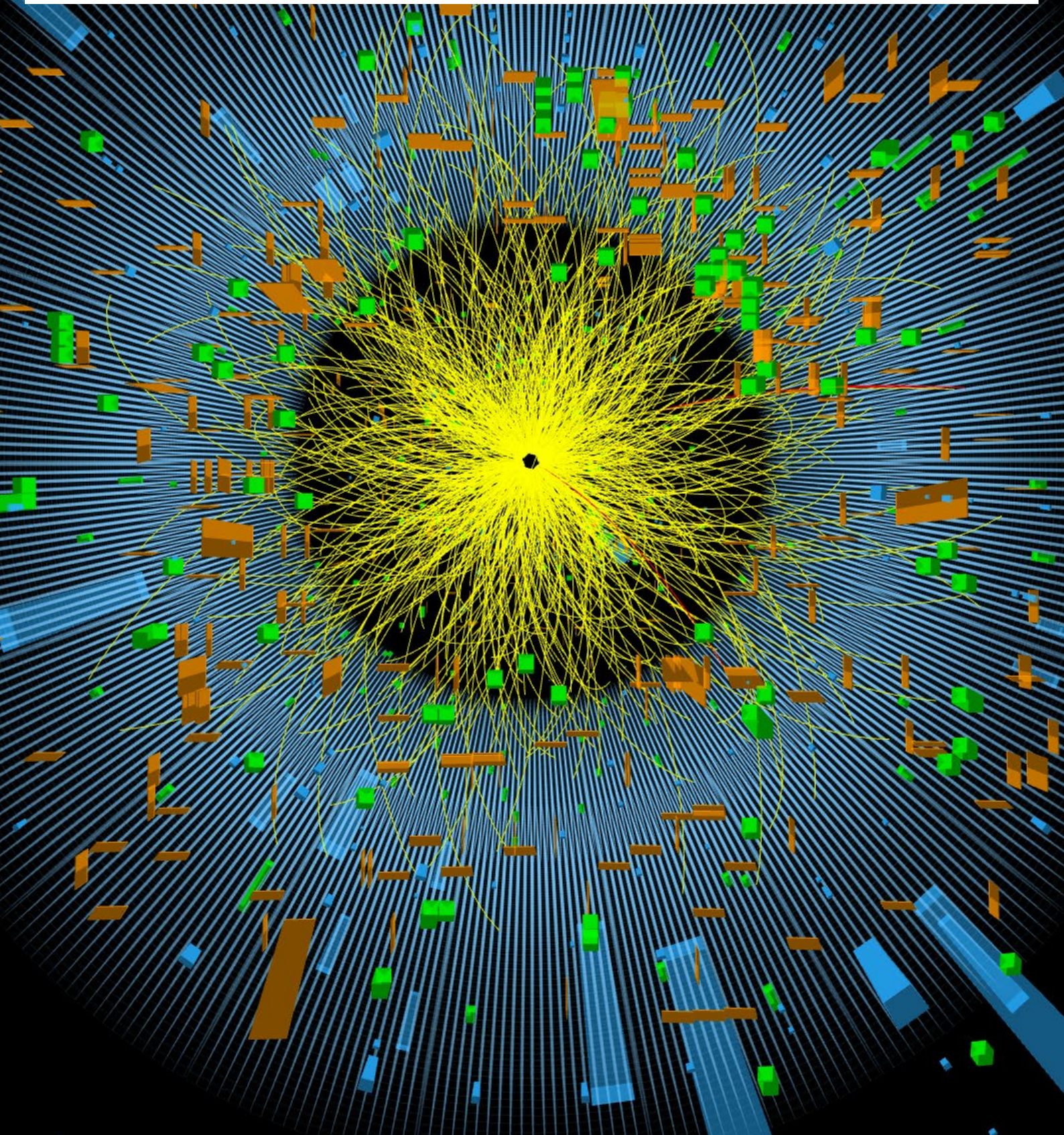




Bundesministerium
für Bildung
und Forschung

Aktionsplan ErUM-Data

Von Big Data zu Smart Data: Digitalisierung in der
naturwissenschaftlichen Grundlagenforschung



Inhalt

Vorwort	2
<hr/>	
Naturwissenschaft im digitalen Wandel: Herausforderungen und Chancen	4
Neue Datentechnologien durch Grundlagenforschung	4
Forschung von morgen.....	6
<hr/>	
Digitale Strukturen für Spitzenforschung – die Ziele von ErUM-Data	8
Wissenschaft vernetzen	8
Digitale Kompetenzen ausbauen	11
Transfer und Kommunikation stärken.....	12
<hr/>	
Übergreifende Lösungen – die Maßnahmen in ErUM-Data	16
Interdisziplinäre Forschungsförderung	17
Qualifizierungsangebote	18
Austausch, Kommunikation und Transfer	19
<hr/>	
Passgenau: die Umsetzung von ErUM-Data	22
Ausgestaltung und Förderprozess	22
Evaluationen zur Erfolgskontrolle	24
<hr/>	
Abkürzungsverzeichnis	26
<hr/>	
Glossar	27
<hr/>	
Impressum	29
<hr/>	

Vorwort

Wir suchen nach den kleinsten Teilchen und nach den größten Strukturen des Universums. Doch was hat das mit dem Alltag der allermeisten Bürgerinnen und Bürger zu tun? Das zeigt beispielsweise das Europäische Forschungszentrum CERN. Hier wurde das World Wide Web entwickelt, das heute unseren Alltag so stark prägt.

Die physikalische Grundlagenforschung stellt sich den großen Fragen unserer Zeit und damit auch ganz besonderen technologischen Anforderungen. Die großen Forschungseinrichtungen erweitern die Grenzen des technisch Machbaren ständig. Und was Wissenschaftlerinnen und Wissenschaftler aus den Themengebieten Teilchen, Materie und Universum heute entwickeln, kann schon sehr bald unser Smartphone schneller machen, unsere Arbeit erleichtern und unsere Freizeit bereichern.

Der Aktionsplan ErUM-Data will diese Entwicklung beschleunigen und steuern. Die Grundlagenforschung profitiert von der schnelleren Digitalisierung. Sie bewältigt riesige Datenmengen und treibt so die digitale Entwicklung zu neuen Lösungen und neuen Erkenntnissen voran. Davon profitieren wir alle.

Gleichzeitig arbeiten die Forscherinnen und Forscher zunehmend fachübergreifend und international zusammen. Denn die globale Gesundheit, der Klimawandel oder demographische Verschiebungen sind komplexe Herausforderungen, die wir nur gemeinsam bewältigen können.

Der einfache Zugang zu Daten und digitalen Werkzeugen ist eine zentrale Säule dieses Aktionsplans. Mit digitalen Wissensspeichern werden Ergebnisse offen zugänglich, aber auch die Forschungsdaten und die Werkzeuge, um diese großen Datenmengen zu analysieren. Mit Open Data und Open Science können sich die Akteure so länderübergreifend viel besser vernetzen.

Der Aktionsplan ErUM-Data schafft dafür den strategischen Rahmen, um das Potenzial der Digitalisierung bei der Erforschung von Teilchen, Materie und Universum voll auszuschöpfen. Auch für die Gesellschaft bieten Digitalisierung und große Datenmengen einen hohen Mehrwert, wenn wir sie verantwortungsvoll und effektiv nutzen.

Ihr Bundesministerium für Bildung und Forschung







Naturwissenschaft im digitalen Wandel: Herausforderungen und Chancen

Daten sind die prägende Ressource des Informationszeitalters. Sie entstehen millionenfach in allen Bereichen des Lebens und Arbeitens. Damit diese Daten wertschöpfend und nutzbringend für die Gesellschaft eingesetzt werden können, sind neue technische Lösungen sowie ein Kulturwandel im Umgang mit Wissen und Daten erforderlich. Das Fundament bilden digitale Technologien.

Datenbasierte Anwendungen sind schon heute allgegenwärtig: Spracherkennungs-Apps, Einparkhilfen und Fitness-Tracker erleichtern unseren Alltag und unterstützen bei einer gesunden Lebensweise. Vernetzte Systeme steuern die Stromversorgung oder sorgen für mehr Sicherheit im Straßenverkehr. Künstliche Intelligenz (KI) bereitet den Weg für smarte Autos, smarte Häuser und smarte Produktion. Maßgeschneiderte Behandlungen einer individualisierten Medizin nutzen den gleichen technologischen Fortschritt wie Frühwarnsysteme für globale Krisen und Katastrophen.

In allen Bereichen nehmen die Menge und Komplexität der Daten stetig zu. Die Anforderungen an Mensch und Technik steigen. Technologien für die Organisation und Verwertung großer Datenmengen werden zum zentralen Motor des Fortschritts: Sie befähigen die Wissenschaften zu Erkenntnissen über unsere Welt, unterstützen die Gesellschaft und fördern die Wirtschaft. Zu der technischen Herausforderung kommen gesell-

schaftliche und ökonomische Aspekte. Die erheblichen Veränderungen durch den digitalen Wandel werfen neue Fragen auf, beispielsweise: Wie lässt sich die steigende Datenflut auch in Zukunft gewinnbringend und verantwortungsvoll für alle nutzen? Welche ethischen Gesichtspunkte müssen dabei bedacht werden und wie bleiben Daten geschützt vor Missbrauch? Wie gelingt es, genügend qualifizierte Datenexpertinnen und -experten auszubilden? Und wie kann die Gesellschaft mit der hohen Geschwindigkeit der digitalen Veränderung Schritt halten? Auf diese und weitere Fragen müssen Wissenschaft, Wirtschaft, Gesellschaft und Politik gemeinsam Antworten finden.

Neue Datentechnologien durch Grundlagenforschung

Die Erforschung von Universum und Materie an Forschungsinfrastrukturen nimmt eine Schlüsselrolle für den digitalen Wandel ein. Als datenintensivster Wissenschaftsbereich ist sie seit jeher mit Big Data, das heißt mit besonders großen und komplexen Datensätzen, konfrontiert. Ihre Forschungsdaten entstehen an Großgeräten, bei Experimenten und aus Simulationen an Hoch- und Höchstleistungsrechenzentren: Astronomische Beobachtungsdaten von Observatorien auf der Erde und im Weltraum helfen Forschenden, unser Universum zu enträtseln. Daten aus unzähligen Kollisionen von Elementarteilchen in großen Teilchen-

Foto auf S. 2/3: Atacama Large Millimeter/submillimeter Array (ALMA) in der chilenischen Atacama-Wüste

Foto oben: Ionenquellen – Startpunkt der GSI/FAIR-Beschleunigeranlage

Forschen mit Big Data

An Forschungsinfrastrukturen gewinnen Wissenschaftlerinnen und Wissenschaftler wertvolle Datensätze für die Erforschung von Universum und Materie. Dabei müssen sie mehrere zentrale Herausforderungen meistern: die **Übertragung riesiger Mengen von Daten in kürzester Zeit**, den **Umgang mit sehr heterogenen, hochkomplexen Datenstrukturen** aus Experimenten an internationalen Großgeräten beziehungsweise Simulationen sowie die **schnelle und direkte Bereitstellung der Daten für Forscherinnen und Forscher weltweit**. Die folgenden Beispiele aus den Themengebieten Teilchen, Materie und Universum illustrieren Chancen und Herausforderungen:



Am weltweit leistungsfähigsten Teilchenbeschleuniger, dem Large Hadron Collider (LHC) an der europäischen Großforschungseinrichtung CERN bei Genf, suchen Forschende nach neuen Teilchen. Dabei produzieren sie heute bereits hundert Mal mehr Rohdaten als sie speichern können. Um die **enorme Datenflut** zu bewältigen, setzen die Forscherinnen und Forscher **Echtzeit-Analysen bei der Datenerfassung** ein. Sie filtern die wichtigsten Daten heraus und archivieren nur diese. Selbst nach dieser Vorauswahl sammeln die Experimente fast 90.000 Terabyte Daten pro Jahr – den Inhalt von etwa 18 Millionen DVDs. Mit der nächsten Ausbaustufe des LHC ab 2027 wird sich diese Zahl nochmals verzehnfachen. Auch die derzeit entstehende Facility for Anti-proton and Ion Research in Europe (FAIR) in Darmstadt wird ab Inbetriebnahme 2027 mit einigen zehntausend Terabyte an Daten pro Jahr und den damit verbundenen Herausforderungen umgehen müssen.



Messungen am ultraschnellen Röntgenlaser European XFEL in Hamburg machen Veränderungen der Materie sichtbar. Dafür sendet der Röntgenlaser bis zu 27.000 Lichtblitze in der Sekunde aus. Bereits bei der Datenerfassung stellen die erforderlichen **hohen Übertragungsraten** eine Herausforderung dar. Jeder Lichtblitz erzeugt ein Streubild – und dieses muss in kürzester Zeit erfasst und weiterverarbeitet werden. Mit Hilfe von extrem schnell arbeitenden Detektoren können die Forscherinnen und Forscher chemische Reaktionen regelrecht filmen und daraus **einzigartige Einblicke** in den Nanokosmos der Materie gewinnen.



Bei Himmelsdurchmusterungen entstehen ständig unzählige Beobachtungsdaten. Schnelle Erstanalysen ermöglichen Astronominnen und Astronomen, auch auf unvorhersehbare, vergängliche Phänomene wie Strahlungsausbrüche von Sternen zu reagieren und zusätzliche Teleskope rechtzeitig auf solche Ereignisse zu richten. **Digitalinfrastrukturen mit geeigneten Speicher- und Zugangslösungen** sind die Voraussetzung, damit Forschungsteams weltweit sowohl **unmittelbar wie auch langfristig mit diesen Daten arbeiten und forschen** können. Viele astronomische Fragestellungen lassen sich nur beantworten, indem die Daten verschiedener Teleskope und Instrumente zusammengeführt werden – teilweise über lange Beobachtungskampagnen hinweg. Da die einzelnen Observatorien immer leistungsfähiger werden, vervielfacht sich die Gesamtmenge und Komplexität der relevanten Daten dadurch immer schneller. Das in der chilenischen Atacama-Wüste entstehende Extremely Large Telescope (ELT) etwa wird als einzelnes Großteleskop bereits in einer einzigen Nacht mehrere Terabyte an Daten sammeln – so viel wie der Informationsgehalt von Milliarden von Buchseiten.





Big Data

Big Data bezeichnet Datenmengen, die so groß und komplex sind, dass herkömmliche Methoden der Datenauswertung nicht ausreichen, um sie zu verarbeiten. Charakteristisch sind die „6 Vs“:

- Volume** Menge, Umfang der Daten
- Velocity** Geschwindigkeit, mit der die Daten erzeugt und übertragen werden
- Variety** Bandbreite der Datentypen und -quellen
- Veracity** Echtheit der Daten
- Value** Wert der Daten
- Validity** Datenqualität



Smart Data

Intelligente Methoden zur Mustererkennung und zur weiterführenden Datenanalyse ziehen gewinnbringende Informationseinheiten aus Big Data. Sie überführen die Daten in eine optimale Form für ihre weitere Verwendung und Langzeitarchivierung. Auf dieser Basis werden der wissenschaftliche Erkenntnisgewinn sowie datengetriebene Innovationen in Wissenschaft und Wirtschaft möglich oder beschleunigt. Neben den aufbereiteten Daten umfasst der Begriff „Smart Data“ auch die zugehörigen Bearbeitungsmethoden und Auswahl-schritte, etwa um die Daten zu filtern, zu clustern oder zu vernetzen.

beschleunigern geben Aufschlüsse zu grundlegenden Fragen der Physik. Daten aus Rechenmodellen bilden physikalische, chemische und biologische Abläufe ab.

Mit steigender Leistungsfähigkeit der Großgeräte und der zugehörigen Digitalinfrastrukturen wächst auch die Menge der Forschungsdaten, die effizient organisiert und gewinnbringend verarbeitet werden müssen. Die Wissenschaftlerinnen und Wissenschaftler stoßen immer wieder an die Grenzen des Machbaren. Sie müssen individuelle Lösungen für den Umgang mit Daten entwickeln oder profitieren von Entwicklungen anderer Forschungsbereiche und der Wirtschaft. Als Erstanwenderin neuer Technologien und Methoden wirkt die Forschung auch als Impulsgeberin für Innovationen und Datendienstleistungen.

Die naturwissenschaftliche Forschungslandschaft ist durch internationale Zusammenarbeit geprägt, Forschungsdaten werden weltweit erzeugt und genutzt. Der Digitalisierungsprozess erfordert daher globales Denken und Handeln. Im Fokus stehen Datenstandards

sowie die Verständigung auf einheitliche Herangehensweisen rund um die Verarbeitung der Daten und den Zugang zu den Daten.

Forschung von morgen

Der Aktionsplan ErUM-Data schafft Voraussetzungen für die naturwissenschaftliche Grundlagenforschung von morgen – an internationalen Forschungsinfrastrukturen, die mit deutscher Beteiligung gebaut und betrieben werden – in Deutschland, Europa und weltweit.

Die neuen technologischen Möglichkeiten bedeuten auch neue Wege im Forschungs- und Erkenntnisprozess: Künstliche Intelligenz und Maschinelles Lernen (ML), kombiniert mit hochleistungsfähigen und vernetzten Digitalinfrastrukturen, erschließen neue Potenziale der Wertschöpfung aus den weltweiten Datenbeständen und beschleunigen den wissenschaftlichen Erkenntnisprozess.

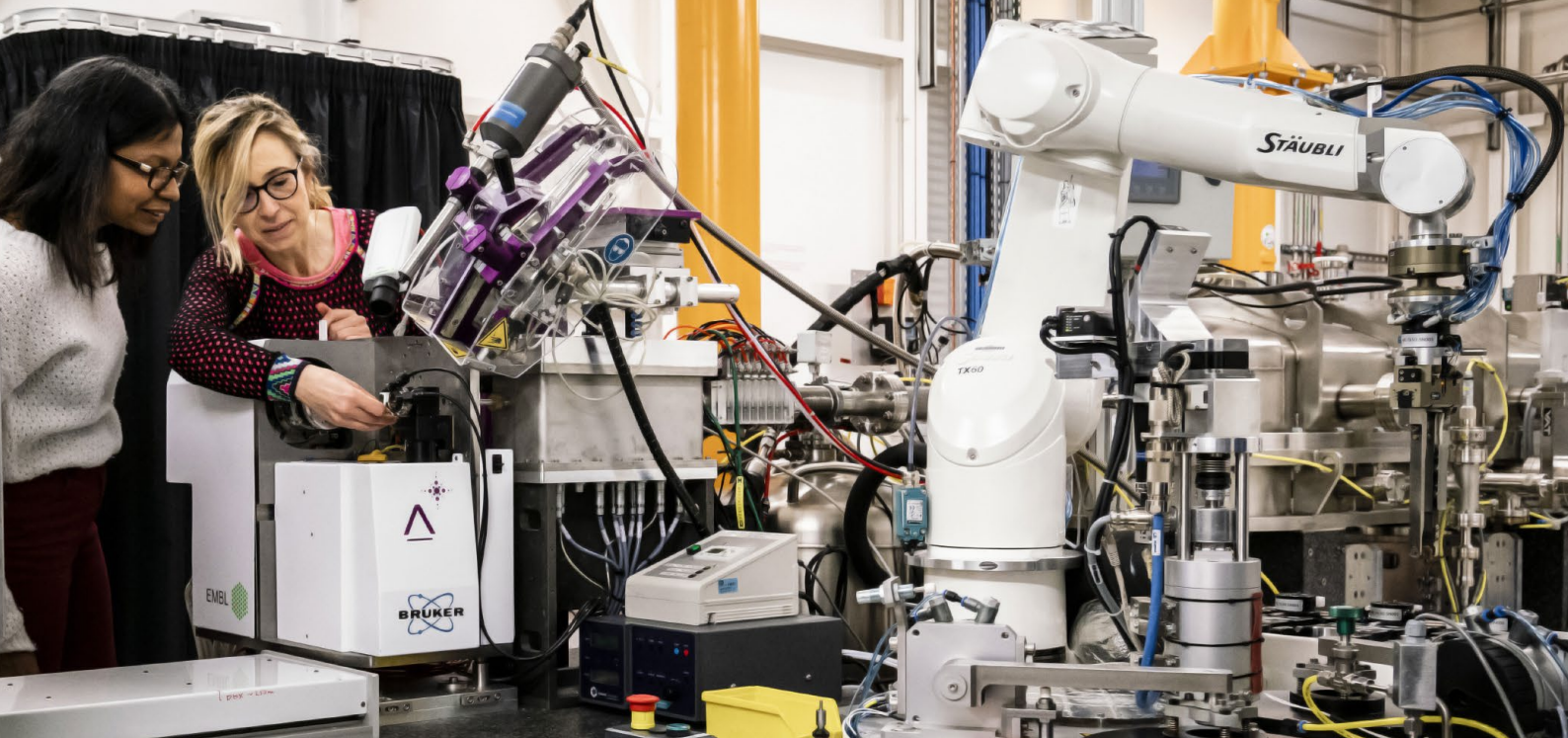
Mit Hilfe immer besserer digitaler Lösungen können Wissenschaftlerinnen und Wissenschaftler in Zukunft fachübergreifend aus Informationsquellen schöpfen, die weit über herkömmliche Publikationen in Fachzeitschriften hinausreichen. Digitale Wissensspeicher werden nicht nur gewonnene Erkenntnisse und Ergebnisse offen zugänglich machen, sondern – wie in einzelnen Projekten bereits Realität – verstärkt auch die zugrundeliegenden Forschungsdaten und Analysewerkzeuge für deren Verarbeitung bereitstellen. Auf diese Weise werden wissenschaftliche Resultate

besser nachvollziehbar. Die Wissensspeicher werden vorhandene Digitalinfrastrukturen in steigendem Maße integrieren, Big und Smart Data-Technologien nutzen und so ein langfristig effektives Forschungsdatenmanagement gewährleisten. Open Data- und Open Science-Konzepte werden einen besseren und vor allem offeneren Zugang zu Daten über Fächer- und Ländergrenzen hinweg ermöglichen. Zugleich eröffnen sie neue Chancen, Wirtschaft und Gesellschaft zu beteiligen.

Digital vernetzte Wissenswelten

Wie die Datennutzung der Zukunft im internationalen Forschungsraum aussehen kann, demonstrieren einige Initiativen mit Modellcharakter aus den ErUM-Themengebieten Teilchen, Materie und Universum. Das BMBF setzt sich dafür ein, solche Vorhaben weiterzuentwickeln und weitere Digitalinfrastrukturen nach diesem Vorbild zu etablieren.

- Das **Worldwide LHC Computing Grid (WLCG)** löst die Herausforderung der großen Datenfluten am LHC-Teilchenbeschleuniger durch weltweit verteilte Rechen- und Speicherkapazitäten sowie neueste Cloud- und Datenbanktechnologien. Es integriert zurzeit rund eine Million Prozessorkerne in über 170 Rechenzentren in 42 Ländern. Über dezentrale Zugangspunkte speisen die über zehntausend Nutzerinnen und Nutzer der vier Hauptexperimente Verarbeitungsaufträge in das System ein, um Daten zu visualisieren und weiterzuverarbeiten. Eine ähnliche föderierte Digitalinfrastruktur ist für das zukünftige internationale Beschleunigerzentrum FAIR vorgesehen. Ziel ist ein möglichst einfacher Zugang zu den Daten und Analyseinstrumenten, der unterschiedliche Software, Speicher- und Rechensysteme integriert. Durch einheitlich gestaltete Schnittstellen sollen Wissenschaftlerinnen und Wissenschaftler eines LHC-Experiments zukünftig problemlos auch Daten eines FAIR-Experiments bearbeiten können – oder umgekehrt.
- Rund 40.000 Wissenschaftlerinnen und Wissenschaftler nutzen Röntgen- und Neutronenquellen in Europa, um Einblicke in die Eigenschaften der Materie zu erhalten. Der **Photon-Neutron Data Science Demonstrator** bietet ihnen eine offene Forschungsumgebung und integrierte Lösung für die Datenspeicherung und -analyse. Die Auswertung eines einzelnen Experiments umfasst dabei in der Regel sowohl die Datenverarbeitung wie auch eine Modellierung und Simulation und erfordert entsprechend große Rechenressourcen. Mit dem Photon-Neutron Data Science Demonstrator sollen Forschungsgruppen zukünftig weltweit die Daten der Röntgen- und Neutronenquellen mithilfe bereitgestellter Software an verteilten Ressourcen auf einer virtuellen Plattform visualisieren und analysieren. Das Projekt dient als Pilot für die European Open Science Cloud (EOSC).
- Die Vision des **Virtual Observatory (VO)** ist es, sämtliches astronomisches Wissen über das Universum digital bereitzustellen: Es soll die Gesamtheit astronomischer Daten und Informationen verknüpfen – als ein stetig wachsendes, einheitliches Ganzes. Dies beinhaltet Beobachtungsdaten in allen zugänglichen Wellenlängenbereichen, die mit modernen Teleskopen erfasst oder in vergangenen Jahrhunderten dokumentiert wurden, sowie Simulationsdaten. Im Virtual Observatory werden digitale Plattformen und Datenarchive in international verteilten Zentren schrittweise zusammengeschlossen. Essenziell für den Projekterfolg sind einheitliche Standards für die Datenstrukturen und ihre Handhabung. Neben Rohdaten bietet das Virtual Observatory funktionale Web-Interfaces sowie Software und Visualisierungsinstrumente, mit denen Nutzerinnen und Nutzer Daten auswählen und weiterverarbeiten können.



Digitale Strukturen für Spitzenforschung – die Ziele von ErUM-Data

Mit ErUM-Data schafft das Bundesministerium für Bildung und Forschung (BMBF) die notwendigen Rahmenbedingungen, um die Herausforderungen der Digitalisierung in der naturwissenschaftlichen Grundlagenforschung schneller und systematischer zu bewältigen – für exzellente Forschung in und aus Deutschland. Der Aktionsplan fördert eine koordinierte Herangehensweise und breite Verständigung. Er ebnet den Weg zu übergeordneten Lösungen. Langfristiges Ziel sind optimale Bedingungen für eine effiziente und finanzierbare Datenverarbeitung und -verwertung.

ErUM-Data fügt sich in das 2017 veröffentlichte Rahmenprogramm „Erforschung von Universum und Materie - ErUM¹“ ein. Das ErUM-Rahmenprogramm setzt forschungspolitische Schwerpunkte für die naturwissenschaftliche Grundlagenforschung an nationalen und internationalen Forschungsinfrastrukturen. Es stellt grundlegende Weichen für wissenschaftliche Spitzenleistungen, Zukunftstechnologien, Innovationskeime und den wissenschaftlichen Nachwuchs von morgen. Der Querschnittsbereich Digitalisierung und der Umgang mit Forschungsdaten sind zentral für alle ErUM-Themengebiete – Teilchen, Materie und Universum.

Mit dem vorliegenden Aktionsplan ErUM-Data definiert das BMBF drei übergeordnete strategische Ziele:

1. Relevante Akteure und Strukturen sollen mit Blick auf die Digitalisierung vernetzt werden;
2. Digitale Kompetenzen in der Forschung sollen ausgebaut werden;
3. Transfer und Kommunikation sollen gestärkt werden – in ErUM und über ErUM hinaus.

Wissenschaft vernetzen



Die digitale Herausforderung für die naturwissenschaftliche Grundlagenforschung kann nur gemeinsam gemeistert werden. ErUM-Forscherinnen und -Forscher müssen sich stärker koordinieren, auf gemeinsame Vorhaben und Standards verständigen und dabei auch weitere Disziplinen einbeziehen, wie Mathematik, Informatik und andere technologiestarke Bereiche.

¹ bmbf.de/de/erforschung-von-universum-und-materie---das-rahmenprogramm-erum-4388.html

Foto oben: Messaufbau an der Röntgenquelle ESRF in Grenoble, Frankreich

Strategischer Rahmen

Die Ziele und Maßnahmen des Aktionsplans ErUM-Data passen sich in das ErUM-Rahmenprogramm sowie in nationale und globale Strategien zur Digitalisierung ein:

Nationales Umfeld

Das BMBF unterstützt die datenbasierte Wissenschaft in Deutschland auf verschiedenen Wegen. Wichtige Initiativen sind etwa der nationale **Rat für Informationsinfrastrukturen** als Beratungsgremium, überregionale nationale **Kompetenzzentren für KI-Forschung** sowie die **Open Access-Strategie**. Mit der 2019 veröffentlichten **Digitalstrategie** unterstreicht das BMBF seine Rolle als zentraler Akteur für die Gestaltung des Digitalen Wandels in Bildung, Wissenschaft, Forschung und Innovation. Ziele sind insbesondere, aus Daten Wissen und Innovationen zu schaffen sowie die technologische Souveränität und wissenschaftliche Vorreiterrolle für Deutschland und Europa zu sichern. Ein zentraler Baustein der Digitalstrategie ist die **Nationale Forschungsdateninfrastruktur (NFDI)**. Die NFDI wird einen föderierten und vernetzten Wissensspeicher aufbauen, der zahlreiche Datenbestände erschließt und systematisiert. Im Bereich des Hoch- und Höchstleistungsrechnens setzt sich das BMBF besonders für Vernetzungsaktivitäten ein, wie sie von der **Gauß-Allianz** unterstützt werden. Sie treibt das Hoch- und Höchstleistungsrechnen als eigenständige strategische Forschungsaktivität voran und hilft Wissenschaftlerinnen und Wissenschaftlern in Deutschland dabei, Kompetenzen und Rechnerstrukturen zu bündeln und zu nutzen. Durch den Zusammenschluss der drei leistungsstärksten deutschen Höchstleistungsrechenzentren zum **Gauss Centre for Supercomputing** bietet Deutschland die zurzeit mit Abstand leistungsfähigste Supercomputing-Infrastruktur Europas für Wissenschaft und Wirtschaft. Seit Ende 2018 wird mit dem **Nationalen Hochleistungsrechnen (NHR)** ergänzend ein Netzwerk von Hochleistungsrechnern errichtet, um das wissenschaftliche Rechnen an Hochschulen zu stärken – bei deutschlandweitem Zugang.

Auf Bundesebene begegnet Deutschland der Digitalisierung durch miteinander verzahnte Strategien. Mit der **Umsetzungsstrategie Digitalisierung** gestaltet die Bundesregierung den digitalen Wandel übergreifend. Eine Mission der **Hightech-Strategie 2025** ist es, durch eine offene Innovationskultur neue Quellen für neues Wissen zu schaffen. Dafür sollen die Möglichkeiten von Open Access Publikationen, Open Data, Open Science und Open Innovation umfassender genutzt werden. Mit der **KI-Strategie** bringt die Bundesregierung Künstliche Intelligenz Made in Germany an die Weltspitze. ErUM-Data ergänzt diese Strategien durch passgenaue Beiträge in Wissenschaft und Forschung.

Weltweite Initiativen

Auch im europäischen Interesse übernimmt das BMBF eine treibende Rolle. Die NFDI sichert den Anschluss an die **European Open Science Cloud (EOSC)**, die zukünftig nationale und europäische Forschungsorganisationen und Dateninfrastrukturen in einem „Web of FAIR data“ verbinden wird. Innerhalb der europäischen **GO FAIR Initiative** wirkt Deutschland auf europaweite Standards im Bereich Forschungsdatenmanagement und auf die übergreifende Nutzung bestehender Infrastrukturen im Sinne der FAIR Data-Prinzipien hin. Mit **GAIA-X** treibt das BMBF gemeinsam mit dem BMWi sowie nationalen und europäischen Partnern den Aufbau eines sicheren und vernetzten europäischen Daten- und Infrastrukturökosystems voran, das den höchsten Ansprüchen an digitale Souveränität genügen soll und in dem Daten und Dienste verfügbar gemacht, zusammengeführt und vertrauensvoll geteilt werden sollen. Die **EuroHPC-Initiative** der Europäischen Union und ihrer Mitgliedstaaten wird europäische und nationale Ressourcen des Hoch- und Höchstleistungsrechnens zu einem europäischen Ökosystem zusammenführen: für eine optimale Supercomputing- und Dateninfrastruktur für Europas Nutzerinnen und Nutzer aus Wissenschaft, Industrie und Gesellschaft. Das BMBF unterstützt die Zielsetzung von EuroHPC aktiv. Aus dem internationalen Forschungsraum werden wichtige Impulse zum Beispiel durch die **Research Data Alliance (RDA)** aufgenommen. Diese Initiative greift in einem Bottom-up-Prozess Herausforderungen im weltweiten Datenmanagement auf und erarbeitet fachübergreifende Empfehlungen, wie sich soziale und technische Barrieren beim Zugang zu Daten abbauen lassen.

Nur so können sie die nötigen Fortschritte im Hinblick auf Digitalinfrastrukturen und Methodenkompetenz erreichen und mit der Dynamik der Digitalisierung Schritt halten. Dies ist Voraussetzung für die effektive und langfristige Wertschöpfung aus Forschungsdaten.

FAIRe Datenstandards für die Praxis

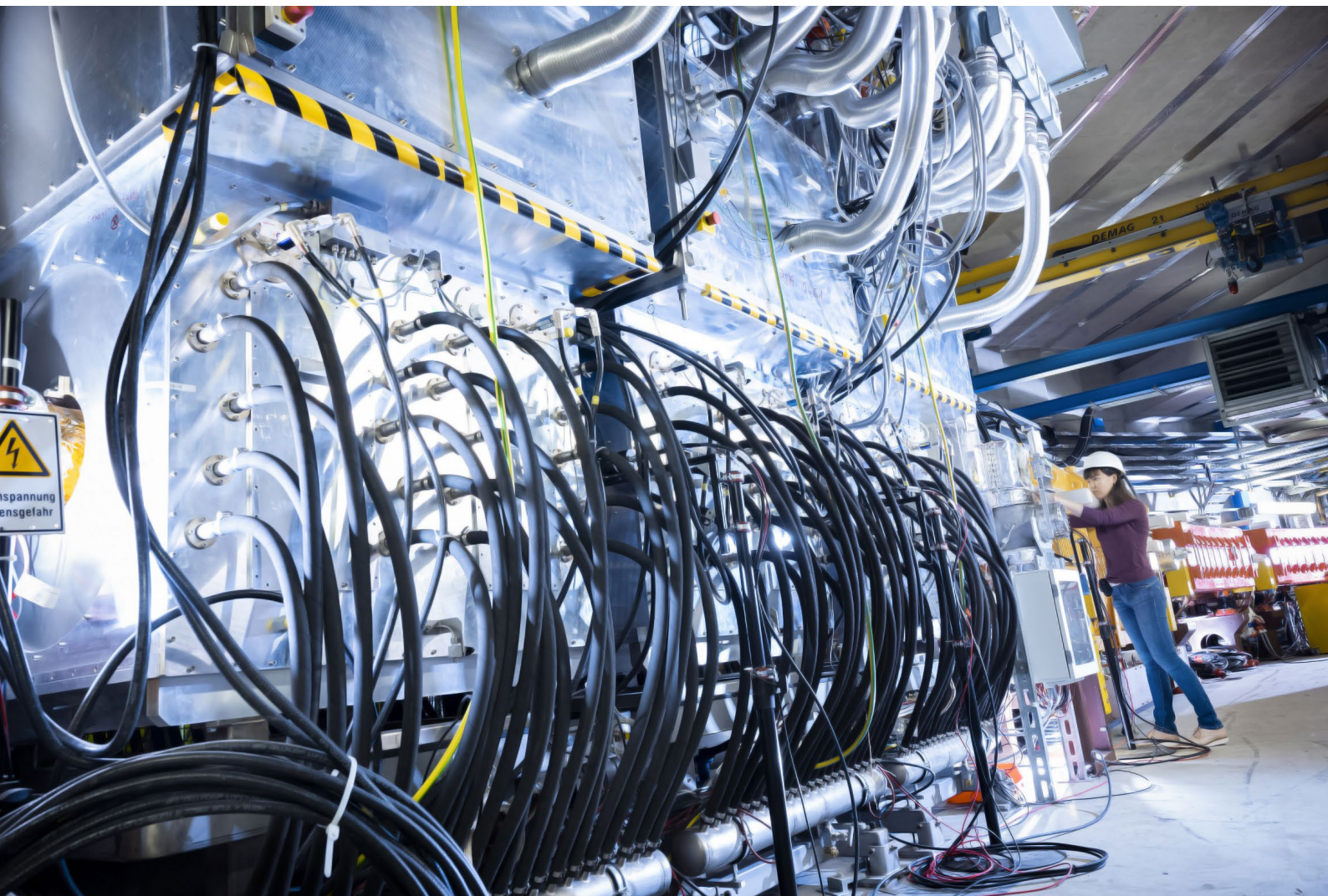
Nur ein transparentes und standardisiertes Datenmanagement über den gesamten Datenlebenszyklus hinweg ermöglicht auf lange Sicht die breite Nutzbarkeit komplexer Datensätze aus Experimenten und Simulationen. Einheitliche, abgestimmte Verfahren sind in den verschiedenen Fachdisziplinen der Naturwissenschaften jedoch noch sehr unterschiedlich weit etabliert. Fach- und länderübergreifende Standards sind notwendig, um sicherzustellen, dass Forschungsdaten langfristig auffindbar, zugänglich, interoperabel und wiederverwendbar (FAIR: findable, accessible, interoperable, reusable) sind. Das BMBF setzt sich dafür ein, in ErUM:

- standardisierte Datenformate für Rohdaten, aufbereitete Daten und Metadaten zu implementieren,
- einheitliche Prozesse, Verfahren und Standards zum Umgang mit Forschungsdaten und ihrer Veröffentlichung nach den FAIR Data-Prinzipien zu definieren,
- Regelungen für den standardisierten, offenen und langfristigen Zugang zu Daten und Publikationen zu schaffen.

Digitalinfrastrukturen verbinden

Besonders an Großgeräten mit sehr vielfältigen Nutzergruppen mangelt es vielfach noch an einer einheitlichen, für alle zugänglichen Digitalinfrastruktur. Derzeit setzen die ErUM-Forscherinnen und -Forscher sehr unterschiedliche technische Lösungen ein, die teilweise nicht mit der Datenmenge mitwachsen oder nur aufwendig an neue Anforderungen angepasst werden

Foto unten: Hochfrequenz-Kavität im Schwerionensynchrotron SIS18



können. Einzelne Forschungseinrichtungen und Großprojekte verfolgen bereits einen förderierten Ansatz. An diesen guten Beispielen orientiert, zielt ErUM-Data darauf ab:

- bestehende Infrastrukturen für Forschungsdaten weiterzuentwickeln, auszubauen und in förderierte Digitalinfrastrukturen zu überführen sowie weitere nach diesem Vorbild zu etablieren,
- die technologischen Hürden zwischen High-Throughput Computing (HTC) und High-Performance Computing (HPC) abzubauen und die Nutzung entsprechender nationaler und europäischer Zentren zu befördern,
- neue Technologien wie Datenspeicherlösungen und Cloud Services einzuführen oder zu adaptieren,
- diese Aktivitäten mit nationalen und internationalen Initiativen wie der NFDI, der EOSC und dem WLCG zu verschränken.

Neue Software und Algorithmen interdisziplinär entwickeln

Das BMBF möchte digitales Know-how der forschungsstarken Disziplinen an der Schnittstelle von naturwissenschaftlicher Grundlagenforschung und Computer- und Datenwissenschaften bündeln. Dies bedeutet:

- gemeinsam leistungsfähige Software und Algorithmen zu entwickeln und zu nutzen, um mit der rasanten Entwicklung der Rechnertechnologie Schritt zu halten und moderne Rechnerarchitekturen für die notwendige Datenverarbeitung und -analyse optimal auszunutzen,
- Methoden aus den Bereichen Künstliche Intelligenz und Maschinelles Lernen verstärkt in der Forschung einzusetzen und weiterzuentwickeln,
- dabei von den Erfahrungen und Kenntnissen relevanter Disziplinen aus ErUM sowie der Mathematik, den Datenwissenschaften und der Informatik – von der theoretischen Methodenentwicklung bis hin zur Erprobung und Anwendung – zu lernen.

Digitale Kompetenzen ausbauen



Kompetenzen im Bereich Datenwissenschaften und der sichere Umgang mit digitalen Methoden und Technologien sind für die Forscherinnen und Forscher in den ErUM-Themengebieten in steigendem Maße entscheidend für den Forschungserfolg.

Mit ErUM-Data schlägt das BMBF starke Brücken zwischen den Natur- und Datenwissenschaften. Ziel ist es, aus ersten Initiativen der Zusammenarbeit langfristig neue Wege der gegenseitigen Unterstützung und des Kompetenzerwerbs sowie neue Karrierepfade zu entwickeln, um den Bedarf an fachübergreifend ausgebildeten Expertinnen und Experten in Wissenschaft und Wirtschaft zu decken.

Nachwuchs- und Weiterqualifizierung in ErUM

In der Praxis müssen viele ErUM-Wissenschaftlerinnen und -Wissenschaftler schon lange mit großen Datenmengen umgehen. Einige bringen bereits wertvolle Big Data-Expertise mit. Sie haben zum Beispiel gezeigt, dass Methoden der Künstlichen Intelligenz und des Maschinellen Lernens an verschiedenen Stellen des Datenlebenszyklus gewinnbringend angewendet werden können. In Zukunft müssen die Forschenden dazu befähigt werden, neue Analysemethoden viel breiter einzusetzen und das gesamte Spektrum an Möglichkeiten für eine effiziente Auswertung der Mess- und Simulationsdaten auszunutzen. Dies bedeutet auch einen Kulturwandel im Umgang mit Wissen und Daten. Das BMBF fördert daher:

- die Ausbildung von ErUM-Nachwuchskräften in neuen digitalen Methoden und Techniken sowie im Datenmanagement bereits ab einer frühen Karrierephase,

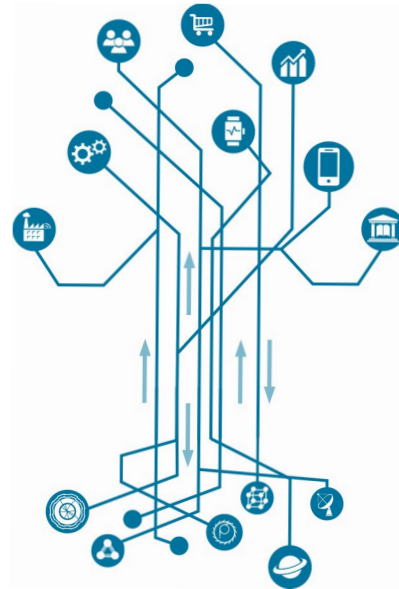
- die systematische fachübergreifende Weiterqualifizierung von etablierten ErUM-Wissenschaftlerinnen und -Wissenschaftlern.

Fächerübergreifend Kompetenzen erwerben

Die Expertise in Forschungsdatenmanagement und digitalen Technologien unterscheidet sich noch stark zwischen den Fächern und Disziplinen innerhalb sowie außerhalb von ErUM. Wissenschaftliche Teilgebiete, die weniger Erfahrung mit Forschungsdaten oder theoretischer Methodenentwicklung haben, können von Fortschritten und Lösungen der Fächer mit einem größeren Erfahrungsschatz im Umgang mit Daten und Softwareentwicklung profitieren. Die übergreifende Weitergabe von Wissen beschleunigt Weiterbildungsprozesse von Fach- und Führungskräften in der Wissenschaft und erweitert damit auch die Fachkräftebasis für die Wirtschaft. Ziel des Aktionsplans ErUM-Data ist es daher:

- die systematische und fächerübergreifende Weitergabe von Wissen anzuregen,
- technologisch starke Disziplinen außerhalb von ErUM in Qualifizierungsprojekte einzubeziehen,
- den intensiven Austausch von Expertise zu befördern.

Transfer und Kommunikation stärken



Deutschland gehört in vielen Bereichen zur Gruppe der weltweiten Innovationsführer. Gerade im Bereich der Digitalisierung sind auch zukünftig rasante Fortschritte mit hohem Innovationspotenzial zu erwarten. Als Entwicklerin oder Erstanwenderin neuer Technologien steht die naturwissenschaftliche Grundlagenforschung am Anfang der Innovationskette. Aus ihren Erkenntnissen und Lösungsansätzen können später marktreife Anwendungen und nützliche Datendienstleistungen für Wirtschaft und Gesellschaft hervorgehen. Dies setzt eine intensive Zusammenarbeit und Kommunikation aller Akteure voraus.

Aus der Elementarteilchenphysik in die Industrie

Für die Wissenschaft entwickelte neuartige Technologien im Bereich von Big Data haben ein großes Innovationspotenzial für die Wirtschaft. So ist beispielsweise das 2008 gegründete Start-up Blue Yonder aus einem BMBF-Forschungsprojekt zu neuronalen Netzen am europäischen Forschungslabor CERN hervorgegangen: Neben dem Firmengründer rekrutiert sich auch ein Großteil des Personals der Firma aus der Elementarteilchenphysik. Blue Yonder bietet Service-Leistungen im Bereich Predictive Analytics an, das heißt Berechnungen, um Muster in Daten aufzudecken und ihre weitere Entwicklung vorherzusagen. Mit Hilfe von Maschinellem Lernen und Künstlicher Intelligenz lassen sich Prognosen für komplexe, datenbasierte Fragestellungen treffen. Beispielsweise können Lebensmittelgeschäfte im Einzelhandel die optimale Vorratshaltung für ihre Waren berechnen lassen. Die Vorhersagen beziehen dabei Abhängigkeiten von weiteren Modellen wie beispielsweise Wetterlagen mit ein. Die Prognosen berücksichtigen auch Nachfragespitzen (beispielsweise Grillgut) und das voraussichtliche Kaufverhalten. Damit werden unnötige Kosten und Abfall vermieden.

Der Erfolg des weiteren Digitalisierungsprozesses in Deutschland hängt zudem auch entscheidend von der Bereitschaft der Gesellschaft ab, den technologischen Wandel zu unterstützen. Dafür ist es wichtig, Forschungsergebnisse verständlich und zielgruppengerecht zu kommunizieren, so dass Vertrauen in neue Technologien entstehen kann. Auch um Nachwuchs für die MINT-Fächer zu gewinnen, spielt die Wissenschaftskommunikation eine zentrale Rolle.

Daten und Wissen teilen

Publikationen, Daten und Informationen sollen auch langfristig für einen möglichst breiten Interessentenkreis sowohl innerhalb als auch außerhalb der Wissenschaft zugänglich sein. So können sie für weitere Forschung genutzt werden und Innovationen anregen. Durch Wissenstransfer können neue Technologien und Methoden aus datenintensiven ErUM-Fächern in die interdisziplinäre Nutzung und in vielfältige Anwendungen überführt werden. Das BMBF fordert:

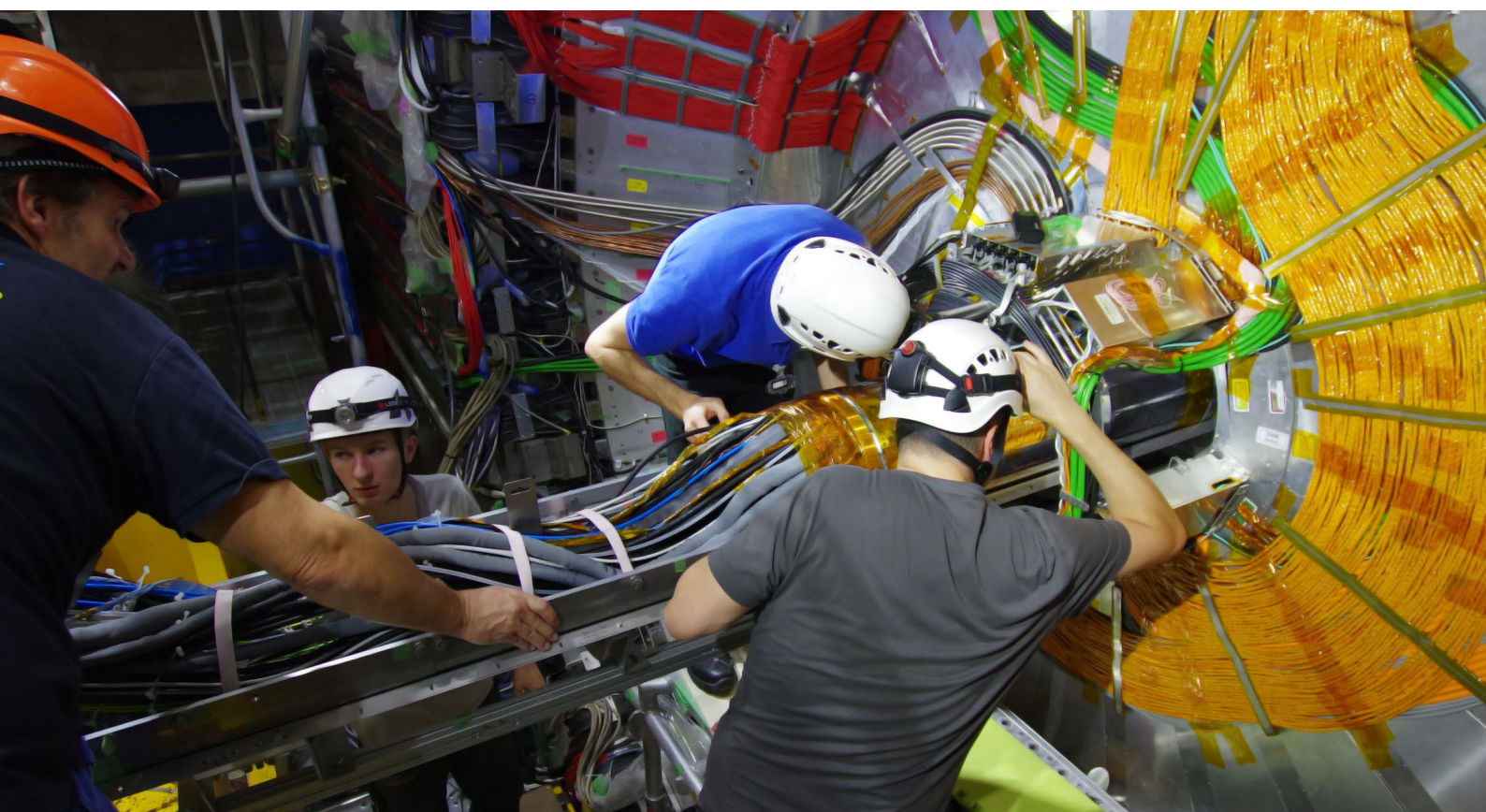
- Open Access-, Open Data- und Open Science-Konzepte umzusetzen und wo nötig anzupassen und weiterzuentwickeln für einen offeneren Austausch von wissenschaftlichen Informationen,
- den themen-, fach- und branchenübergreifenden Wissenstransfer zu Methoden von Künstlicher Intelligenz und Maschinellem Lernen auszubauen, um aktuelle Entwicklungen von Software und Algorithmen überall schnell nutzen zu können.

Innovations- und Technologietransfer steigern

Während Fächer wie Informatik oder Mathematik als Methodenentwickler wirken, beschleunigt die naturwissenschaftliche Grundlagenforschung den Innovationsprozess durch ihre anspruchsvollen Anwendungsfelder. Vielfach werden dort neue Methoden erstmals erprobt. Dies befördert bahnbrechende Entwicklungen bis hin zu Sprunginnovationen. Das BMBF setzt sich dafür ein, den fächerübergreifenden Austausch insbesondere auch mit der Wirtschaft zu neuen Methoden und Technologien zu stärken, um:

- vielversprechende Ansätze aus der Forschung in die kommerzielle Anwendung zu übertragen,
- Entwicklungen aus der Wirtschaft an Forschungsdaten zu testen oder bereits ausgereifte, kommerziell eingesetzte Methoden der Wirtschaft auf die Forschung zu übertragen,

Foto unten: Belle-II-Experiment am japanischen Forschungszentrum für Teilchenphysik KEK



- durch effizienten Wissenstransfer aktuelle Entwicklungen schnell nutzen zu können,
- Lücken in der Innovationskette zu schließen, angefangen von den ErUM-Fächern als Erstanwender neuer Methoden und Technologien bis hin zur wirtschaftlichen Verwertung.

Wissenschaft erfolgreich kommunizieren

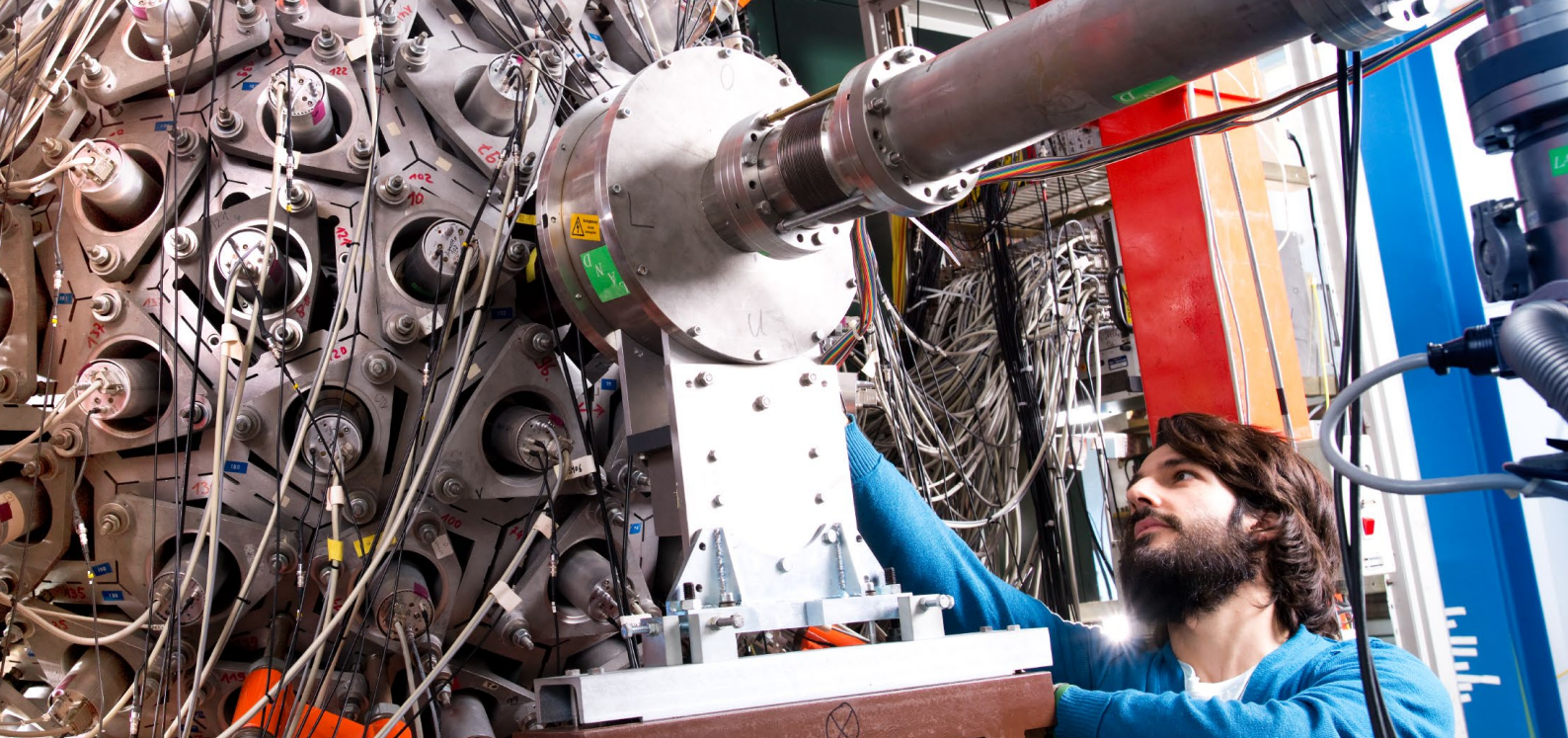
Die Digitalisierung erweitert die Möglichkeiten der Wissenschaftskommunikation erheblich und stellt gleichzeitig veränderte Anforderungen an die Kommunikation. Über digitale Medien findet ein direkter Austausch zwischen Wissenschaft, Gesellschaft, Wirtschaft und Politik statt. Open Science erleichtert dabei den Zugang zu Wissen für alle. Durch Citizen Science eröffnet die Digitalisierung den Bürgerinnen und Bürgern neue Wege, an Forschung teilzuhaben. Dies bereichert umgekehrt auch die Forschung in ErUM. Das BMBF verfolgt die Ziele:

- ErUM-Forschungsergebnisse verständlich und transparent an Zielgruppen außerhalb der Wissenschaft zu vermitteln, diese zu informieren und Interesse an Forschungsthemen und Technologien zu wecken – besonders auch bei Schülerinnen und Schülern sowie Studieninteressierten,
- Bürgerinnen und Bürger frühzeitig und in geeigneter Weise in aktuelle Forschung in ErUM einzubeziehen und einen gesellschaftlichen Diskurs zu ermöglichen.

Teilchenkollisionen, Galaxien und schwarze Löcher: Citizen Science für Universum und Materie

Algorithmen zur Mustererkennung in großen Datenmengen lernen durch Training, Daten effektiver auszuwerten. Die Forschung kann dabei von der Beteiligung interessierter Bürgerinnen und Bürger profitieren. Das Citizen Science-Projekt **Radio Galaxy Zoo** (zooniverse.org) unterstützt beispielsweise die astronomische Datenauswertung, indem Besucherinnen und Besucher der Webseite in Aufnahmen des europäischen Radioteleskops LOFAR Galaxien aufspüren oder Materie rund um Schwarze Löcher identifizieren. Diese Einschätzung per Auge verbessert die computergestützte Klassifizierung. Citizen Science Projekte wie **LHC@home** (lhathome.web.cern.ch) schaffen eine digitale Infrastruktur für die Auswertung von Teilchenkollisionen aus den gebündelten freien Rechenkapazitäten privater Computer der Teilnehmenden.





Übergreifende Lösungen – die Maßnahmen in ErUM-Data

Mit ErUM-Data definiert das BMBF wichtige nächste Schritte für die Digitalisierung der naturwissenschaftlichen Grundlagenforschung an Großgeräten – hin zu mehr interdisziplinärer Zusammenarbeit, gezielter Aus- und Weiterbildung sowie mehr Raum für Austausch und Wissenstransfer. Der Aktionsplan definiert Maßnahmen in drei ErUM-übergreifenden Themenfeldern:

- Forschungsdatenmanagement,
- Föderierte Digitalinfrastrukturen,
- Software und Algorithmen.

Damit die Wissenschaftlerinnen und Wissenschaftler an naturwissenschaftlichen Forschungseinrichtungen weiterhin Spitzenforschung leisten können, müssen Fortschritte in allen drei übergreifenden Themenfeldern erzielt werden.

Forschungsdatenmanagement

Um Daten optimal zu nutzen, müssen diese über ihren gesamten Lebenszyklus effizient organisiert werden – von der Aufzeichnung, über die Übertragung, Auswahl, Verarbeitung und Analyse der Daten bis hin zu ihrer Archivierung und Bereitstellung für Dritte. Grundvoraussetzung sind zukunftsfähige Verarbeitungs- und Nutzungskonzepte, die den Zugriff auf

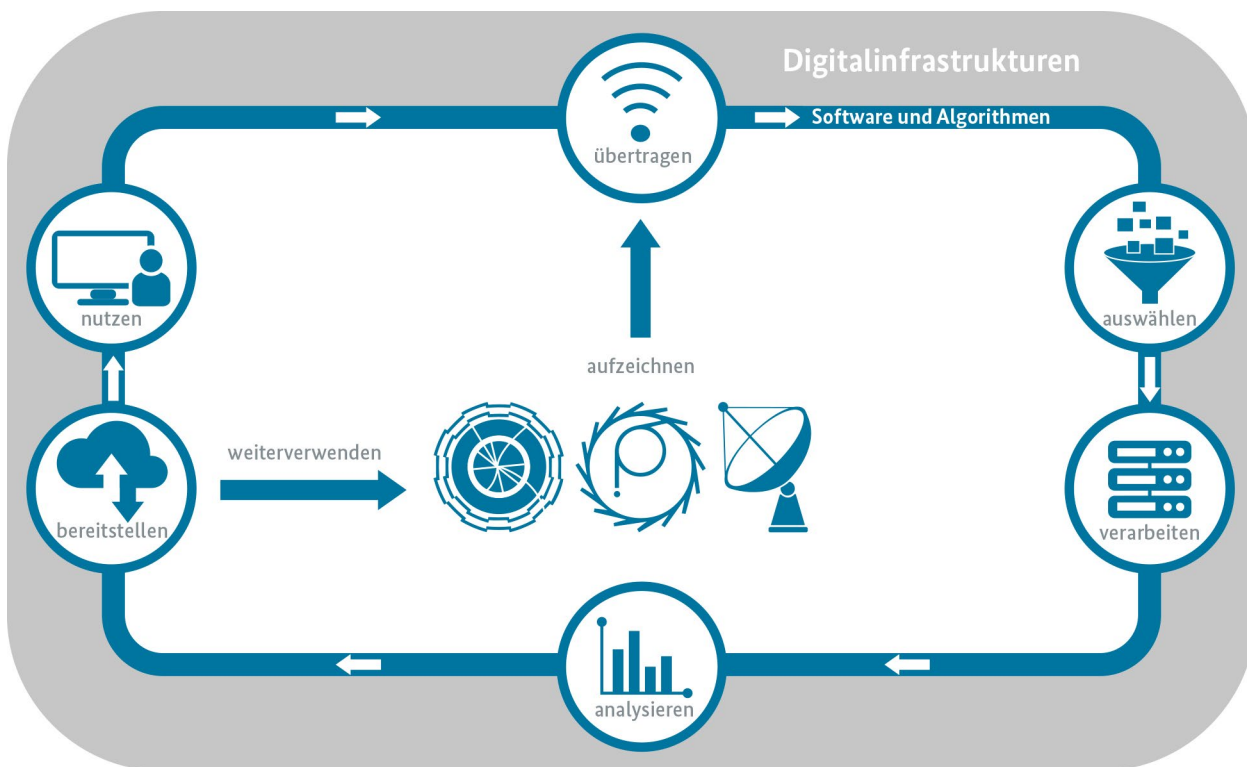
Forschungs- und Metadaten langfristig regeln. Diese Konzepte müssen die Randbedingungen unterschiedlichster, internationaler Großgeräte und Fachdisziplinen berücksichtigen und sich an FAIR Data-Prinzipien orientieren.

Föderierte Digitalinfrastrukturen

Skalierbare Lösungen sind unerlässlich für den Umgang mit der riesigen Menge an existierenden und zukünftigen Forschungsdaten. Das teils explosive Wachstum der Forschungsdaten an Großgeräten fordert zudem vielfach Weiterentwicklungen und Anpassungen, denen nur föderierte Lösungen gerecht werden können. Dazu werden individuelle Rechen-, Speicher- und Netzwerkkapazitäten zu föderierten hochleistungsfähigen Digitalinfrastrukturen zusammengeschlossen. Diese Kapazitäten müssen vorausschauend geplant werden. Neben Spitzenlasten steigen auch die Anforderungen an die Langzeitarchivierung und Nachnutzung der Daten. Hand in Hand mit der Weiterentwicklung und dem Ausbau von Hardware geht die Entwicklung geeigneter Software, wie Datenbank- und Cloud-Technologien, um die Infrastrukturen effizient zu betreiben und zu nutzen.

Software und Algorithmen

Neue Software und Algorithmen befördern einerseits die Umsetzung und Effizienz technischer Lösungen



Forschungsdatenmanagement wird über den gesamten Datenlebenszyklus hinweg betrieben – vom Aufzeichnen der Daten bis hin zu ihrer Auswertung und langfristigen Bereitstellung. Föderierte Digitalinfrastrukturen sowie neue Software und Algorithmen sind notwendig, um Daten besser zu organisieren und neue Erkenntnisse zu gewinnen.

für Digitalinfrastrukturen und sind andererseits für die wissenschaftliche Weiterverarbeitung und den Erkenntnisgewinn aus großen und komplexen Datenmengen erforderlich. Big und Smart Data-Methoden ermöglichen, Daten mit zuvor unerreichbarer Genauigkeit und Geschwindigkeit auszuwerten. Auf Künstlicher Intelligenz und Maschinellem Lernen basierende Verarbeitungs-, Analyse- und Visualisierungstechniken erkennen selbstständig Muster in den Daten und leiten Zusammenhänge ab. Die ErUM-Forscherinnen und -Forscher werden zunehmend intelligente und skalierbare Algorithmen nutzen und müssen diese weiterentwickeln, um neue Schlüsse aus den Forschungsdaten zu ziehen.

Das BMBF fördert zielgerichtet Maßnahmen in diesen drei ErUM-Data-Themenfeldern. Dazu gehören die interdisziplinäre Forschungsförderung, Qualifizierungsangebote sowie Maßnahmen im Bereich Austausch, Kommunikation und Transfer.

Interdisziplinäre Forschungsförderung

Das BMBF unterstützt mit ErUM-Data die interdisziplinäre Zusammenarbeit von Hochschulen, Forschungs- und Rechenzentren sowie der Wirtschaft. Die entsprechenden Vorhaben adressieren eines oder mehrere der drei ErUM-Data-Themenfelder.

Forschungs- und Entwicklungsvorhaben

Die Förderung richtet sich im Wesentlichen an interdisziplinäre Forschungs- und Entwicklungsvorhaben in Kooperation mit den ErUM-Forschungsinfrastrukturen. Zum einen sollen themengebieteübergreifende Vorhaben innerhalb von ErUM über die Bereiche Teilchen, Materie und Universum hinweg unterstützt werden. Zum anderen rückt ErUM-Data Projekte in den Fokus, die zusätzliche Expertise von Fächern außerhalb von ErUM – wie beispielsweise der Mathematik, der Informatik oder den Datenwissenschaften – einbeziehen.

Die Zusammenarbeit kann durch die Forschungsförderung in unterschiedlicher Form angebahnt und gestärkt werden, etwa durch Kooperationen im Management von Forschungsdaten, durch die gemeinschaftliche Weiterentwicklung und den Zusammenschluss bestehender ErUM-Digitalinfrastrukturen oder durch interdisziplinäre Ansätze im Bereich der Algorithmen- und Softwareentwicklung.

Wissens- und Technologietransfer

Mit der Forschungsförderung in ErUM-Data nimmt das BMBF auch solche Vorhaben in den Blick, die Partner aus der Wirtschaft aktiv einbeziehen und Innovationsketten durch frühzeitigen Austausch anstoßen können. Diese Aktivitäten sollen Potentiale erschließen und idealerweise zu beständigen Partnerschaften zwischen Wissenschaft und Wirtschaft führen. Dazu gehören Forschungs- und Entwicklungsarbeiten in ErUM, die:

- den Transfer der Ergebnisse in die Anwendung vorsehen,
- vorhandenes Know-how aus der Wirtschaft zu innovativen Methoden und Technologien nutzen.

Bündelung von Kompetenzen

Der interdisziplinäre Charakter der Ziele und Maßnahmen von ErUM-Data erfordert eine regelmäßige organisationsübergreifende Zusammenarbeit zu drängenden Themen und Anlässen. Dies soll durch thematisch fokussierte Vernetzungsplattformen, die sogenannten ErUM-Data-Impulsräume, unterstützt und befördert werden. Die Impulsräume sollen unterschiedliche Akteure und deren Expertise zusammenbringen und Fragestellungen aus übergeordneter Perspektive beleuchten. Sie helfen, Aktivitäten zu bündeln und auszurichten, und befördern die schnelle Entwicklung und Umsetzung von Ideen durch ein koordiniertes Vorgehen. Als Kompetenzschnittstellen unterstützen sie Bottom-up-Aktivitäten. Wo möglich, sollen die Impulsräume an etablierte BMBF-Kompetenzzentren anknüpfen und von diesen profitieren.

Qualifizierungsangebote

Komplementäre Qualifizierungsangebote sollen Forscherinnen und Forscher in unterschiedlichen Karrierephasen ansprechen – von Studierenden bis zu erfahrenen Forscherinnen und Forschern. Das BMBF legt großen Wert darauf, dass neue Qualifizierungsprogramme mit bestehenden Angeboten regional und überregional abgestimmt und koordiniert werden. Die Maßnahmen müssen für die ErUM-Themengebiete Teilchen, Materie und Universum konzipiert sein. Falls bereits ähnliche Aktivitäten existieren, sollen diese passend ergänzt werden.

Ausbildungsmaßnahmen

Mit ErUM-Data unterstützt das BMBF Aktivitäten, die insbesondere Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftler möglichst früh ansprechen, an digitale Themen heranzuführen und gewinnbringend qualifizieren. Diese Maßnahmen sollen sowohl in die Tiefe als auch in die Breite gehen. Sie sollen bewährte wie moderne Formate aufgreifen, wie beispielsweise:

- Research Schools zu ErUM-übergreifenden Digitalthemen,
- Summer & Winter Schools mit hohem Praxisbezug,
- Angebote wie Workshops und Massive Open Online Courses (MOOC).

Spezifische Weiterbildungsmaßnahmen

Etablierte ErUM-Wissenschaftlerinnen und -Wissenschaftler benötigen vielfach digitales Spezialwissen für ihre datenbasierten Forschungsarbeiten und für die Weiterentwicklung digitaler Infrastrukturen der Experimente und Großgeräte. Oft ähneln sich die Herausforderungen an den Forschungsinfrastrukturen methodisch, so dass koordinierte Qualifizierungsmaßnahmen im Vordergrund stehen. Diese Maßnahmen sollen außerdem mit Angeboten verzahnt werden, die über ErUM hinausreichen: etwa mit Initiativen der Research Data Alliance, der Gauß-Allianz (künftig auch des NHR), des EuroHPC und des Netzwerks der Kompetenzzentren für KI-Forschung. Mögliche Maßnahmen sind beispielsweise Workshops, Trainings oder Hackathons.

Karrierperspektiven

Mithilfe von Nachwuchsförderprogrammen für ErUM-Data will das BMBF Anreize für die Hochschulen und Forschungszentren schaffen, digital qualifiziertes Personal langfristig an den Einrichtungen zu halten. Insbesondere soll Personal in fachübergreifenden Querschnittsbereichen der ErUM-Data-Themenfelder unterstützt werden, wie etwa im Bereich der Entwicklung von übergreifend nutzbaren Algorithmen und Rechenmodellen. Darüber hinaus setzt sich das BMBF dafür ein, dass sich der wissenschaftliche Nachwuchs

mit seinen wissenschaftlich-technischen Kompetenzen für die weitere Karriere in Wissenschaft und Wirtschaft profilieren kann.

Austausch, Kommunikation und Transfer

Die enge Abstimmung aller relevanten Akteure ist essenziell, damit gemeinsame Lösungen für die Digitalisierung identifiziert werden können. Ein strukturierter Vernetzungsprozess soll Interessenvertreterinnen und Interessenvertreter innerhalb und außerhalb von ErUM beteiligen und die Themengebiete Teilchen, Materie und Universum näher zusammenführen. Für eine erfolgreiche Gestaltung und die notwendige Akzeptanz muss dieser Prozess aus den ErUM-Fächern heraus entstehen: Bottom-up, unterstützt durch gezielte Förderung.

Zusammenarbeit und Transfer

Die Umsetzung der strategischen Ziele und der zugehörigen Maßnahmen in ErUM-Data erfordert ein abgestimmtes und professionelles Vorgehen. Das BMBF plant deshalb, frühzeitig eine zentrale Vernetzungs- und Transferstelle zu fördern, im Folgenden als ErUM-Data-Hub bezeichnet. Der ErUM-Data-Hub soll maßnahmenübergreifend angelegt werden und Akteure aus allen ErUM-Themengebieten und ErUM-Data-Themenfeldern einbeziehen. Er soll Ideen aus der Wissenschaftsgemeinschaft aufgreifen und bei deren Umsetzung unterstützen und koordinieren. Dies betrifft insbesondere Aktivitäten der Aus- und Weiterbildung, des Wissens- und Innovationstransfers und der Öffentlichkeitsarbeit. Der Hub soll sich in all diesen Bereichen eng mit den ErUM-Data-Vorhaben abstimmen. Auch soll er beim Transfer von entwickelten Lösungen in neue Anwendungsfelder helfen und Partner über ErUM hinaus gewinnen, beispielsweise zu den Themen Energie, Klima, Nachhaltigkeit oder Gesundheit.

Foto unten: Ingenieurin und IT-Spezialist mit Server-Racks im Modern Data Center



Austausch zu Digitalthemen

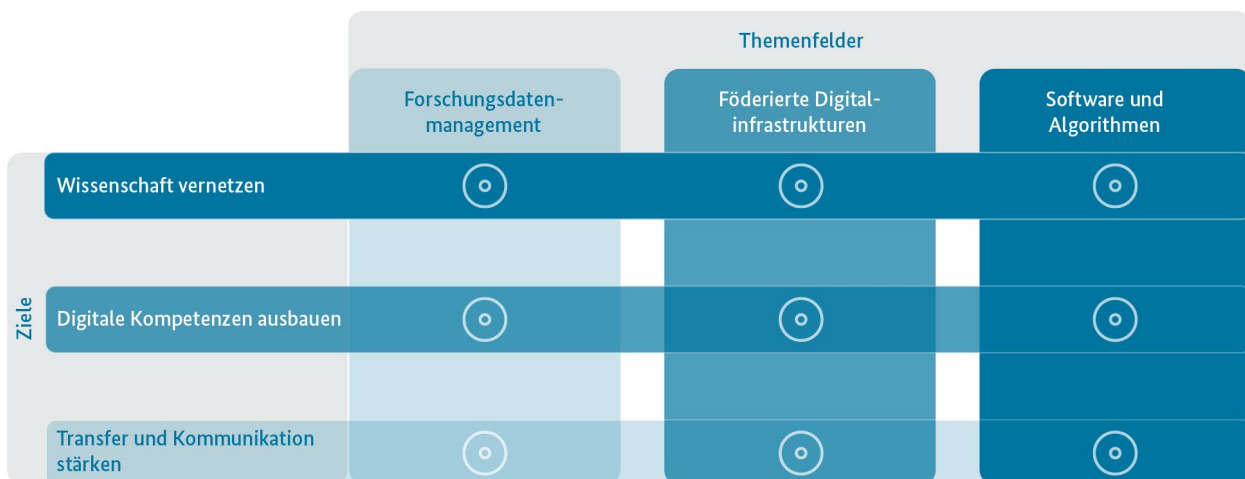
Mit der Förderung in ErUM-Data stärkt das BMBF Aktivitäten, die explizit den Austausch und Dialog zu digitalen Querschnittsthemen im Blick haben. Geeignete Formate sollen sicherstellen, dass Expertenwissen weitergegeben werden kann und aktuelle Entwicklungen und Bedarfe der Digitalisierung rechtzeitig erkannt und berücksichtigt werden. Mögliche Formate sind:

- Meetings, Konferenzen und Workshops,
- Gastaufenthalte: Personalaustausch zwischen Hochschulen, Forschungsinfrastrukturen, Rechenzentren sowie der Wirtschaft,
- Austauschplattformen für Transferaktivitäten zwischen Wissenschaft und Wirtschaft.

Dialog mit der Öffentlichkeit

Digitale Themen müssen noch breiter und professioneller kommuniziert werden: aus der Wissenschaft heraus in die Gesellschaft. Erprobte Formate der Wissenschaftskommunikation sollen genutzt werden, um Zielgruppen gezielt anzusprechen. Daneben können neue, innovative Formate entwickelt werden, die den besonderen Themen und dem Charakter der Digitalisierung gerecht werden. Mögliche Maßnahmen umfassen:

- Aktivitäten der Wissenschaftskommunikation und Öffentlichkeitsarbeit, die an geförderte ErUM-Data-Vorhaben anknüpfen,
- gemeinsame Informationsangebote durch ErUM-Data-Verbünde,
- Teilhabe- und Dialogformate sowie Projekte im Bereich der Bürgerwissenschaften.



Maßnahmen in ErUM-Data

Das BMBF legt die übergreifenden Ziele von ErUM-Data für die drei Themenfelder Forschungsdatenmanagement, Föderierte Digitalinfrastrukturen sowie Software und Algorithmen fest. Es fördert aufeinander abgestimmte, passgenaue Maßnahmen in allen Bereichen.

Foto auf S. 21: Laserlicht erzeugende Undulatoren (gelb) des European XFEL



Gewicht: 8,0t
Weight: 8,0t

KFF

KFF

KFF



Passgenau: die Umsetzung von ErUM-Data

ErUM-Data hält geeignete Förderinstrumente und Maßnahmen bereit, die den Digitalisierungsprozess in der naturwissenschaftlichen Grundlagenforschung voranbringen. Diese ergänzen weitere Förderaktivitäten des BMBF passgenau und erlauben, flexibel auf wissenschaftliche und gesellschaftliche Entwicklungen in der Digitalisierung zu reagieren. Sie sind eng mit relevanten Maßnahmen und Initiativen der Bundesregierung und der Europäischen Union verzahnt.

Investitionen in Wissenschaft, Forschung und Lehre sind eine gesellschaftliche Aufgabe, die Bund und Länder gemeinsam tragen. Hierfür nutzen sie mit der institutionellen Förderung, der Projektförderung sowie regelmäßigen Beiträgen zum Haushalt internationaler Forschungseinrichtungen drei zentrale Förderinstrumente. Für ErUM sind diese Instrumente als Teil des übergeordneten Rahmenprogramms eng aufeinander abgestimmt. Im vorliegenden Aktionsplan ErUM-Data wird insbesondere die Projektförderung als flexibles, zielgerichtetes und maßnahmenorientiertes Instrument eingesetzt, das eine kurz- und mittelfristige Gestaltung des Digitalisierungsprozesses erlaubt.

Die im Aktionsplan definierten Fördermaßnahmen adressieren aktuelle und künftige Herausforderungen des digitalen Wandels. Sie sind verzahnt mit der Forschungsförderung des 2018 veröffentlichten

Aktionsplans ErUM-Pro², greifen allerdings spezielle Anforderungen im Zusammenhang mit der Digitalisierung auf.

Ausgestaltung und Förderprozess

Das BMBF sieht für den Aktionsplan ErUM-Data eine Laufzeit von 2020 bis 2030 vor. Die Dauer der einzelnen Maßnahmen richtet sich bedarfsgerecht an den förderrechtlichen Möglichkeiten aus. Unter Vorbehalt verfügbarer Haushaltsmittel sind für die Fördermaßnahmen insgesamt bis zu 120 Millionen Euro vorgesehen.

Mit dem Rahmenprogramm ErUM hat das BMBF den Prisma-Prozess als zentralen Strategieprozess eingeführt. Die Prisma-Instrumente wurden bereits bei der Vorbereitung des vorliegenden Aktionsplans eingesetzt, um die wissenschaftlichen Akteure der ErUM-Themengebiete Teilchen, Materie und Universum sowie Vertreterinnen und Vertreter aus Gesellschaft und Wirtschaft gezielt einzubeziehen. Neben Prisma-Trialogen und dem Prisma-Forum fand ein Strategieworkshop zu Herausforderungen und Chancen der Digitalisierung in der naturwissenschaftlichen Grundlagenforschung statt. In diesem Zuge wurden Bedarfe aus Wissenschaft, Gesellschaft und Wirtschaft identifiziert. In Folge dieses Workshops entstand auch ein Abschlusspapier aus Sicht der ErUM-Community³.

² [bmbf.de/de/neuer-aktionsplan-erum-pro-ernetzt-hochschulen-und-forschungsinfrastrukturen-7339.html](https://www.bmbf.de/de/neuer-aktionsplan-erum-pro-ernetzt-hochschulen-und-forschungsinfrastrukturen-7339.html)

³ Challenges and Opportunities of Digital Transformation in Fundamental Research on Universe and Matter, Recommendations of the ErUM Committees (2019)

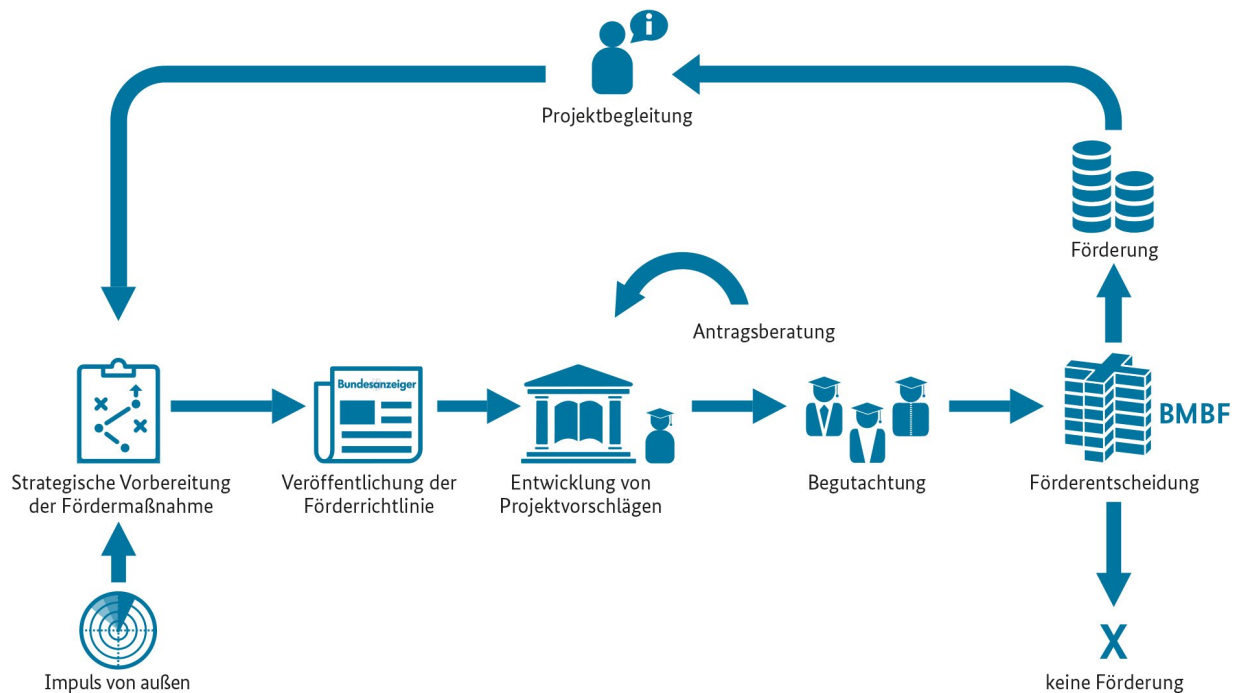
Zusätzlich wurden im Rahmen einer internationalen Vergleichsstudie wichtige Impulse aus anderen forschungsstarken Ländern aufgegriffen.

Für den Aktionsplan ErUM-Data hat das BMBF die Impulse und Bedarfe im Hinblick auf ihre forschungspolitische Bedeutung bewertet und daraus die hier vorgestellten konkreten Ziele und Maßnahmen entwickelt.

In einem ersten Schritt zur Umsetzung wird das BMBF prioritäre Schwerpunkte in themenspezifischen Förderrichtlinien aufgreifen, die mittelfristig in thematisch breitere Förderrichtlinien überführt werden sollen. Gemäß der aktuellen Prioritäten des BMBF und des festgestellten Bedarfs der ErUM-Community an neuen Techniken und Methoden sollen dabei zunächst Forschungs- und Entwicklungsarbeiten aus dem ErUM-Data-Themenfeld Software und Algorithmen im Bereich Künstliche Intelligenz und Maschinelles Lernen gefördert werden. Daran anknüpfend sind Fördermaßnahmen zum Themenfeld Forschungsdatenmanagement vorgesehen, gefolgt von Maßnahmen zu weiteren prioritären Bereichen aller drei Themenfelder.

Wichtige übergreifende Aufgaben wie die ErUM-spezifische Aus- und Weiterbildung, der Wissens- und Innovationstransfer und die Öffentlichkeitsarbeit sollen von Beginn an über alle Maßnahmen hinweg durch den ErUM-Data-Hub unterstützt werden. Zur Mitte der Laufzeit ist eine dedizierte Förderrichtlinie zur Zusammenarbeit mit der Wirtschaft vorgesehen. Es ist geplant, für die weitere Ausgestaltung auch Statusseminare zu nutzen, in deren Rahmen die geförderten Aktivitäten vorgestellt und evaluiert werden und die darüber hinaus dem Wissens- und Erfahrungsaustausch der geförderten Projekte in den unterschiedlichen Themenfeldern dienen sollen.

Das BMBF wird bei der Umsetzung von ErUM-Data, ähnlich dem Vorgehen im Aktionsplan ErUM-Pro, im Vorfeld von Fördermaßnahmen in der Regel Prisma-Strategiegespräche durchführen. Dieses Instrument erlaubt es, Anregungen für die Ausgestaltung der geplanten Fördermaßnahmen von relevanten Akteuren sowohl aus den ErUM-Themengebieten als auch aus den ErUM-Data-Themenfeldern und den angrenzenden Wissenschaftsbereichen zu erhalten.



Die Förderrichtlinien werden in einem strategischen Prozess erarbeitet, in den auch Impulse von außen aufgenommen werden. Die strategische Vorbereitung beinhaltet in der Regel auch ein Prisma-Strategiegespräch. Geförderte Vorhaben begleitet ein Projektträger über deren gesamte Laufzeit.

Evaluationen zur Erfolgskontrolle

Forschung und Entwicklung sind dynamische Prozesse. Sich wandelnde ökonomische und gesellschaftliche Rahmenbedingungen beeinflussen diese zusätzlich. Daher ist es wichtig, den Erfolg der Fördermaßnahmen zu überprüfen. In den Förderrichtlinien werden dazu Zielvorgaben formuliert. Nach diesen werden

die Umsetzung der förderpolitischen Programmziele, die Wirksamkeit der Förderung sowie die Einbindung und Zusammenarbeit der Akteure bewertet – und zwar sowohl begleitend wie auch nach Laufzeitende des Aktionsplans. Die gewonnenen Erkenntnisse bilden die Grundlage, um den Aktionsplan und die Fördermaßnahmen für die Zukunft weiterzuentwickeln und an sich ändernde Rahmenbedingungen anzupassen.

Die Instrumente des Prisma-Prozesses

Der Prisma-Prozess ist der zentrale Strategieprozess von ErUM: Hochschulen, Wissenschaftsorganisationen und der internationale Forschungsraum – dargestellt durch die drei Seiten eines Prismas – tragen als wissenschaftliche Akteure die Erforschung von Teilchen, Materie und Universum. Mit dem Prisma-Prozess verknüpft das BMBF die individuellen Strategien der Akteure und leitet – unter Berücksichtigung der Interessen von Gesellschaft und Wirtschaft – eine kohärente Gesamtstrategie ab.

- Mit dem **Prisma-Radar** beobachtet das BMBF mit Unterstützung des zuständigen Projektträgers kontinuierlich die wissenschaftlichen und gesellschaftlichen Entwicklungen im Kontext des Rahmenprogramms ErUM. Daraus ergeben sich Impulse für die Prisma-Foren, -Trialoge, -Strategiegespräche und -Konferenzen.
- Das **Prisma-Forum** ist das Beratungsgremium des Ministeriums zur strategischen Ausrichtung von ErUM. Es ist mit hochrangigen Persönlichkeiten mit unterschiedlichen Blickwinkeln auf das Rahmenprogramm besetzt.
- Die **Prisma-Trialoge** dienen dazu, die Interessenvertreterinnen und -vertreter aus Hochschulen, Wissenschaftsorganisationen und dem internationalen Forschungsraum zusammenzuführen. Vertreterinnen und Vertreter dieser Seiten erarbeiten Handlungsempfehlungen und abgestimmte Konzepte zu einem Querschnittsthema oder Handlungsfeld von ErUM.
- **Prisma-Strategiegespräche** dienen der Ausgestaltung von Fördermaßnahmen. Das Ministerium erhält in den Gesprächen Anregungen zu Förderbekanntmachungen von Fachleuten aus den ErUM-Themengebieten und benachbarter Bereiche.
- Die **Prisma-Konferenzen** sind Diskussionsplattformen für übergeordnete Themen aus Wissenschaft und Gesellschaft und erschließen neue Impulse und Blickwinkel für ErUM. Sie gewährleisten eine breite Kommunikation mit der Öffentlichkeit.



Radar



Forum



Trialoge



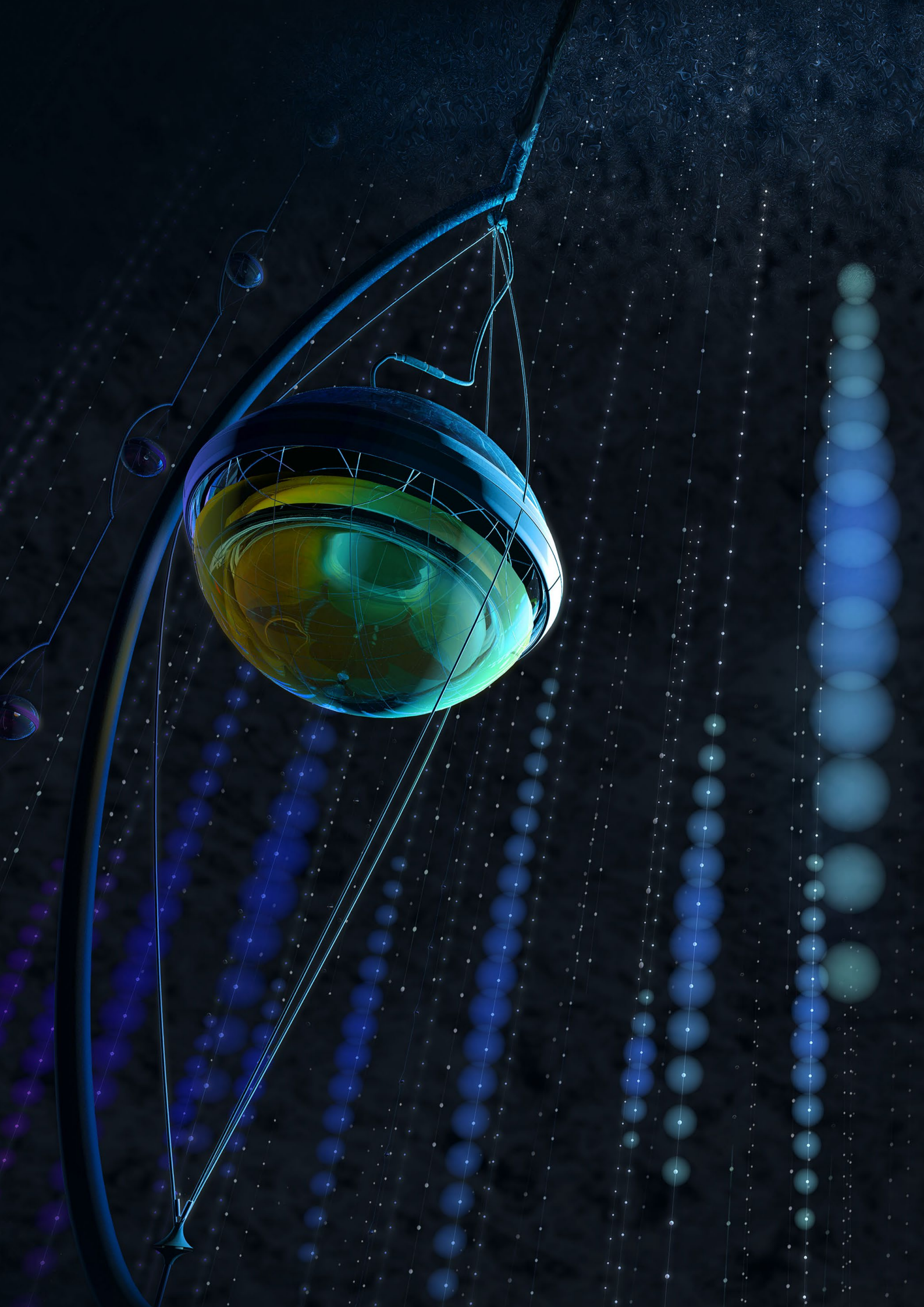
Strategiegespräche



Konferenzen



Strategie



Abkürzungsverzeichnis

ALMA

Atacama Large Millimeter/submillimeter Array

BMBF

Bundesministerium für Bildung und Forschung

CERN

Europäisches Kernforschungszentrum in Genf, ursprünglich: Conseil européen pour la recherche nucléaire

ELT

Extremely Large Telescope

EOSC

European Open Science Cloud

ErUM

Erforschung von Universum und Materie

ESO

European Southern Observatory

European XFEL

European X-Ray Free-Electron Laser Facility

FAIR (Daten)

Findable, Accessible, Interoperable, Reusable Data (auffindbare, zugängliche, interoperable und wiederverwendbare Daten)

FAIR (Beschleunigeranlage)

Facility for Antiproton and Ion Research in Europe

HPC

High-Performance Computing

HTC

High-Throughput Computing

KI

Künstliche Intelligenz

LHC

Large Hadron Collider

MOOC

Massive Open Online Courses

MINT

Mathematik, Informatik, Naturwissenschaften, Technik

ML

Maschinelles Lernen

NFDI

Nationale Forschungsdateninfrastruktur

NHR

Nationales Hochleistungsrechnen

RDA

Research Data Alliance

VO

Virtual Observatory (Virtuelles Observatorium)

WLCG

Worldwide LHC Computing Grid

Glossar

Algorithmen

Eine definierte Abfolge von Anweisungen, um Informationen schrittweise (durch IT-Systeme) zu verarbeiten.

Big Data

Riesige, komplexe Datenmengen, die besondere Methoden der Datenauswertung erfordern. Charakterisiert durch die „6 Vs“: Volume, Velocity, Variety, Veracity, Value, Validity.

Cloud-Technologien

Technologien, mit denen Daten und Informationen dezentral gespeichert, abgerufen und weiterverarbeitet werden können.

Digitalinfrastruktur

Geeignete Hard- und Software-Ausstattung für die Verarbeitung und das Management von (Forschungs-) Daten – vom Erfassen digitaler Daten über ihre Analyse bis zur Langzeitbereitstellung.

ErUM

2017 veröffentlichtes Rahmenprogramm zur Erforschung von Universum und Materie.

ErUM-Pro

2018 veröffentlichter Aktionsplan im ErUM-Rahmenprogramm.

EuroHPC

Europäische Initiative zum Hoch- und Höchstleistungsrechnen.

Experiment

Wissenschaftliche Untersuchung, um eine Hypothese empirisch zu bestätigen oder zu verwerfen. An den ErUM-Forschungsinfrastrukturen werden aufwendige Großexperimente und astronomische Beobachtungen zu grundlegenden naturwissenschaftlichen Fragestellungen durchgeführt.

Forschungsinfrastruktur

Für Forschungsaktivitäten notwendige Ausstattung. In ErUM: Großgerät inklusive aller Infrastruktur vor Ort wie Gebäude, Experimente sowie eigene Digitalinfrastruktur (Hard- und Software).

FAIRe Daten

Daten, die auffindbar, zugänglich, interoperabel und wiederverwendbar sind (Findable, Accessible, Interoperable und Reusable) – gilt auch für zugeordnete Algorithmen, Werkzeuge, Protokolle und Services.

Forschungsdaten

Jegliche digitale und analoge Daten mit wissenschaftlichem Zweck, von Rohdaten aus Experimenten oder Aufzeichnungen bis hin zu Simulationen und Forschungsergebnissen.

GAIA-X

Europäisches Cloud-Projekt zur Entwicklung einer vernetzten Dateninfrastruktur als Wiege eines vitalen, europäischen Ökosystems.

GO FAIR

Europäische Initiative für einheitliche und faire Datenstandards.

HPC (Hoch-/Höchstleistungsrechnen)

Mit Hilfe großer Rechenleistung werden in kürzester Zeit eng gekoppelte Aufträge parallel ausgeführt, beispielsweise zur Berechnung komplexer Modelle.

HTC (Hoch-/Höchstleistungsrechnen)

Viele Rechenressourcen bewältigen über einen längeren Zeitraum Stück für Stück eine Aufgabe, die sich in passende Einzelprobleme zerlegen lässt.

Künstliche Intelligenz (KI)

Ein Teilgebiet der Informatik mit dem Ziel, Maschinen zu einer „intelligenten“ Problemlösung zu befähigen, ein Problem zu abstrahieren und den besten Lösungsweg zu finden.

Maschinelles Lernen (ML)

Schlüsseltechnologie der Künstlichen Intelligenz. ML-Methoden erkennen Muster in Daten. Sie können aus Erfahrung lernen und Daten analysieren, die zu komplex sind, um sie analytisch zu beschreiben.

Open Access Publikation

Eine Veröffentlichung, die im Volltext für alle Menschen frei zugänglich ist: kostenlos und möglichst ohne technische und rechtliche Barrieren.

Open Data

Daten, die – bei Nennung des Urhebers – von allen Menschen frei genutzt, weiterverwendet und geteilt werden können.

Open Science

Wissenschaftspraxis, die den Gedanken des offenen Umgangs mit Daten und Informationen konsequent umsetzt und den Forschungsprozess und die Wissenschaftslandschaft entsprechend gestaltet.

Simulation

Eine Simulation charakterisiert physikalische, chemische oder biologische Systeme durch mathematische Vorschriften, die sich computer-gestützt anwenden lassen, um das Verhalten des Systems vorherzusagen.

Smart Data

Daten in eine Form gebracht, in der diese optimal weiterverwendet werden können, sowie zugehörige Methoden und Technologien.

Ansprechpartnerin

DESY Projektträger

Dr. Charlotte Rothfuchs-Engels

Tel.: 040 8998 5023

E-Mail: charlotte.rothfuchs@desy.de

Impressum

Herausgeber

Bundesministerium
für Bildung und Forschung (BMBF)
Referat Materie und Universum
53170 Bonn

Diese Publikation wird als Fachinformation des Bundesministeriums für Bildung und Forschung kostenlos herausgegeben. Sie ist nicht zum Verkauf bestimmt und darf nicht zur Wahlwerbung politischer Parteien oder Gruppen eingesetzt werden.

Bestellungen

schriftlich an
Publikationsversand der Bundesregierung
Postfach 48 10 09
18132 Rostock
E-Mail: publikationen@bundesregierung.de
Internet: bmbf.de
oder per
Tel.: 030 18 272 272 1
Fax: 030 18 10 272 272 1

Stand

Januar 2022

Text

BMBF
DESY Projektträger, Hamburg
science³ Wissenschaftskommunikation GbR, Berlin

Gestaltung und Redaktion

DESY Projektträger, Hamburg
science³ Wissenschaftskommunikation GbR, Berlin

Druck

BMBF

Illustrationen und Symbole

DESY Projektträger, Hamburg

Bildnachweise

Titel: CERN/Thomas Mc Cauley
S. 2: BMBF/Laurence Chaperon
S. 2/3: ESO/Y. Beletsky (LCO)
S. 4: GSI Helmholtzzentrum für Schwerionenforschung/J. Hosan
S. 8: ESRF/Stef Candé
S. 10: GSI Helmholtzzentrum für Schwerionenforschung/J. Hosan
S. 13: KEK
S. 15: ESO/MPE/Marc Schartmann
S. 16: HA Hessen Agentur/J. Hosan
S. 19: iStock/gorodenkoff
S. 21: European XFEL/Heiner Mueller-Elsner
S. 22: CERN/Maximilien Brice
S. 25: NSF/Ice Cube/Nicolle R. Fuller

