

# Forschungsdaten

MICHAEL DIEPENBROEK, GEROLD WEFER

## Forschungsdateninfrastrukturen in den Bio- und Geowissenschaften

Mit der wachsenden Bedeutung der Verfügbarkeit wissenschaftlicher Daten sind weltweit viele Initiativen, Programme und Projekte zur Schaffung entsprechender Infrastrukturen entstanden. Für die Erdsystemforschung sind besonders erwähnenswert das »Global Earth Observing System of Systems« (GEOSS) und das »World Data System« (WDS) des »International Council for Science« (ICSU). Neben der Entwicklung der notwendigen technischen und organisatorischen Strukturen und Standards hat vor allem die Schaffung von Anreizen zur Archivierung und Publikation von Daten Einfluss auf die nachhaltige Verbesserung der Verfügbarkeit und Qualität. In Deutschland gibt es eine Reihe von etablierten und an das WDS angeschlossenen Datenzentren, die sich 2003 zu einem nationalen Archivnetzwerk zusammengeschlossen haben. Das Konsortium plant für die nächsten Jahre den Aufbau eines Kompetenzzentrums für Forschungsdaten der Erdsystemforschung. Kernstück des auf die Planungen für das ICSU WDS abgestimmten Vorhabens ist die organisatorische und technische Zusammenarbeit zwischen Datenarchiven, Verlagen und Bibliotheken sowie der Aufbau einer entsprechenden webbasierten personalisierbaren Serviceplattform.

Along with the growing importance of the availability of scientific data, many initiatives, programs and projects for the creation of a corresponding infrastructure have arisen around the world. Especially notable in the field of earth system research are the »Global Earth Observing System of Systems« (GEOSS) and the »World Data System« (WDS) of the International Council for Science (ICSU). Along with the development of the necessary technical and organizational structures and standards, it was the creation of incentives for archiving and publishing data which influenced the on-going improvement of accessibility and quality. In Germany a number of data centers connected with the WDS formed a national archival network in 2003. In the coming years this consortium plans to create a competence center for research results in earth system research. At the core of this endeavour, which is in alignment with the plans of the ICSU and WDS, is the organizational and technical collaboration between data archives, publishers and libraries, as well as the development of a corresponding web-based, personalizable service platform.

Die Verfügbarkeit wissenschaftlicher Daten ist einer der Schlüsselfaktoren für die weitere Entwicklung der empirischen Wissenschaften. Bereits 2000 wurde vom Wissenschaftsrat (WR) angemerkt: »Künftig wird es darum gehen, technisch grundsätzlich überall und jederzeit verfügbares Wissen so zu organisieren, dass es überschaubar bleibt und zugleich die bewährten Qualitätsstandards der Wissenschaft gewahrt werden. Hieraus ergeben sich Fragen für die künftige Rolle von Bibliotheken und vergleichbaren Einrichtungen«. <sup>1</sup> Zehn Jahre später hebt der WR erneut die Bedeutung von Informationsinfrastrukturen als konstitutiven Teil des Wissenschaftssystems hervor: »Die Gewinnung neuer wissenschaftlicher Erkenntnisse wäre ohne sie [die Informationsinfrastrukturen] nicht möglich.« <sup>2</sup>

Dieser Aufsatz gibt einen kurzen Überblick über erfolgreiche Initiativen, neuere Konzepte sowie vorhandene Infrastrukturen zur Langzeitarchivierung und

Publikation von bio- und geowissenschaftlichen Daten. Ein besonderer Fokus liegt dabei auf den nationalen Aktivitäten und Einrichtungen.

### GLOBALER ÜBERBLICK

Im globalen Kontext gibt es eine Vielzahl von Initiativen zur Verbesserung der Verfügbarkeit wissenschaftlicher Daten. Prominent sind das »Global Earth Observing System of Systems« (GEOSS) der »Group on Earth Observation« (GEO), <sup>3</sup> das »World Meteorological Organization Information System« (WIS) <sup>4</sup> oder das »World Data System« (WDS) <sup>5</sup> des »International Council for Science« (ICSU). Auf der europäischen Ebene hat sich vor allem das »European Strategy Forum on Research Infrastructures« (ESFRI) mit der Roadmap für eine Paneuropäische Forschungsinfrastruktur engagiert. <sup>6</sup> Darüber hinaus gibt es eine Reihe von fachlich orientierten ebenfalls globalen Infrastrukturen. Dazu gehören u. a. die »International Oceanographic Data and Information Exchange« (IODE) <sup>7</sup> der »Intergovernmental Oceanographic Commission« (IOC) der UNESCO und die »Global Biodiversity Information Facility« (GBIF). <sup>8</sup>

GEOSS vertritt ein ganzheitliches System von der Datenproduktion, Assimilation und Auswertung bis hin zu Entscheidungsprozessen, die wiederum Art und Umfang der Forschungsaktivitäten beeinflussen, konzentriert sich primär auf Erdbeobachtungssysteme und ist ein starker Motor für die Entwicklung von Geodaten-Standards. Wegen fehlender finanzieller Ausstattung ist GEOSS allerdings auf die Rekrutierung von Kapazitäten in den jeweiligen Bereichen angewiesen. Der europäische Beitrag zu GEOSS ist das »Global Monitoring for Environment and Security« (GMES) <sup>9</sup> als Initiative der Europäischen Kommission und der »European Space Agencies« (ESA). GMES geht voraussichtlich 2015 in die operationelle Phase. Einen weiteren europäischen Baustein liefert die »Infrastructure for Spatial Information in the European Community« (INSPIRE), die sich vorrangig mit der Standardisierung von Geoinformation befasst. <sup>10</sup>

Das WIS ist eine Neuauflage des »Global Telecommunication System« (GTS). WIS ist ein technisch anspruchsvolles interoperables System für den Austausch meteorologischer Daten und dient vornehmlich der globalen Produktion von Wettervorhersagen. Daneben sind auch Einrichtungen beteiligt, die kom-



Michael Diepenbroek

Foto: MARUM, Univ. Bremen



Gerold Wefer

Foto: V. Diekamp, MARUM, Univ. Bremen

Verfügbarkeit wissenschaftlicher Daten

plementäre Datenwelten bereitstellen. WIS-Einrichtungen werden von der WMO zertifiziert und akkreditiert. Die Fertigstellung ist für 2013/14 geplant.

Das ICSU WDS ist aus der Revision des seit mehr als 50 Jahren bestehenden Netzwerks des ICSU »World Data Center« (WDC) und der »Federation of Astronomical and Geophysical Data Analysis Services« (FAGS)<sup>11</sup> hervorgegangen. Es ist eine »Non-Governmental Organisation« (NGO) mit der primären Zielsetzung der Schaffung eines föderierten globalen Systems zur Erfassung, Archivierung und Publikation von Daten aus den Geo-, Bio- und Sozialwissenschaften. Wesentliche Aufgabe des WDS ist dabei die organisatorische und technische Vernetzung untereinander sowie mit Bibliotheken und Wissenschaftsverlagen. Ein kohärentes Gesamtsystem bestehend aus Datenarchiven und traditionellen Informationsanbietern im Wissenschaftsbereich bietet sowohl den organisatorisch-technischen als auch den fachlichen Hintergrund, um wissenschaftliche Datenwelten mit ihrer Heterogenität und Dynamik auf anerkannten Wegen zu publizieren und langfristig zu erhalten. Für Wissenschaftler eröffnet ein solches System erheblich bessere Möglichkeiten der Datenrecherche und mehr Sicherheit bei der Verwendung von Daten Dritter. Vor allem aber offeriert die Publikation von Daten eine weitere qualifizierte Möglichkeit, die eigene wissenschaftliche Arbeit zu dokumentieren und damit auch die entsprechende Anerkennung zu erfahren (Publikationen = Wissenschaftswährung).

Das WDS möchte der Publikation wissenschaftlicher Daten als neuer Veröffentlichungsform zu einem breiteren Durchbruch verhelfen. Im Vordergrund stehen dabei Fragen der Qualitätssicherung, die von den Datenarchiven trotz der zumeist guten Einbettung in wissenschaftliche Einrichtungen nicht allein bearbeitet werden können. Mehr als 100 Einrichtungen und Organisationen weltweit haben sich aktuell für das WDS beworben. Das auf dem »Open Archival Information System« (OAIS)<sup>12</sup> und dem im nestor-Projekt der Deutschen Nationalbibliothek erarbeiteten Kriterienkatalog<sup>13</sup> beruhende Zertifizierungsverfahren gewährleistet dabei die Langfristverfügbarkeit und Integrität archivierter Daten.

## NATIONALE EINRICHTUNGEN

Mit dem Ziel der nachhaltigen Verbesserung von Datenverfügbarkeit und -qualität in den Geo- und Biowissenschaften haben sich die in Deutschland ansässigen ICSU-Weltdatenzentren 2003 zu einem nationalen Cluster zusammengeschlossen. Dies sind das AWI/MARUM mit PANGAEA® – Data Publisher for Earth & Environmental Science (vormals WDC-MARE),

das Deutsche Klimarechenzentrum (DKRZ) mit dem WDC for Climate (WDCC) und die Deutsche Luft- und Raumfahrt (DLR) mit dem WDC for Remote Sensing (WDC-RSAT). Dem Verbund haben sich im Weiteren das Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum (GFZ) und die Technische Informationsbibliothek (TIB) angeschlossen. Alle Einrichtungen werden im Anhang kurz beschrieben.

Das Konsortium ist bereits sehr gut mit internationalen Forschungsorganisationen vernetzt, z. B. mit dem »Intergovernmental Panel on Climate Change« (IPCC)<sup>14</sup>, dem »Integrated Ocean Drilling Program« (IODP)<sup>15</sup>, dem »International Continental Drilling Program« (ICDP)<sup>16</sup> oder dem »Global Earth Observations System of Systems« (GEOSS). Gemeinsam wird seit Jahren an der Weiterentwicklung von Standards sowie an Organisations- und Infrastrukturen für Forschungsdaten gearbeitet. So wurde in den letzten Jahren mit Förderung der DFG ein Registrierungsdienst für wissenschaftliche Daten aufgebaut (DOI registry), welcher mittlerweile von einem internationalen Bibliotheken-Konsortium getragen und weltweit genutzt wird (DataCite<sup>17</sup>). Ein besonderer Meilenstein war auch die Implementierung eines gemeinsamen Web-Dienstes mit dem Verlag Elsevier zur Cross-Referenzierung von Publikationen mit supplementären Datensätzen.<sup>18</sup>

WDC sind von der ICSU mandatierte Langzeiteinrichtungen. Darüber hinaus wird von allen Einrichtungen ein WMO-Mandat gehalten. Für den Grundbetrieb sind dauerhafte Finanzierungen in den tragenden Einrichtungen vorhanden. Aufwandsabhängige Ausgaben werden über Drittmittel aus Forschungsprojekten abgedeckt. Alle Einrichtungen sind interoperabel entsprechend internationaler Standards (W3C<sup>19</sup> / ISO<sup>20</sup> / OGC<sup>21</sup> u. a.). Sie sind maßgebend für die Organisation und Implementierung der zentralen Dateninfrastruktur und des Datenportals für das WDS sowie national und global in vielen Informationsportalen vertreten.

## AUSBLICK

Entsprechend der konzeptionellen Vorlage des ICSU WDS ist ein Kompetenzzentrum als Bindeglied zwischen wissenschaftlichen Einrichtungen, Verlagen, Bibliotheken und dem bestehenden Archivnetzwerk für Daten aus Erd- und Umweltforschung geplant (Abb.). Das Vorhaben wird von der DFG gefördert. Konkret sollen nachhaltige und verlässliche Wege zur Publikation wissenschaftlicher Daten geschaffen werden, die den Qualitätsstandards wissenschaftlichen Publizierens entsprechen. Dazu ist beabsichtigt, eine aus der Sicht der Nutzer einheitliche Instanz zu

Vernetzung untereinander  
und mit Bibliotheken und  
Verlagen

Publikationen als  
»Wissenschaftswährung«

Cluster der deutschen  
Geo- und Biowissenschaften

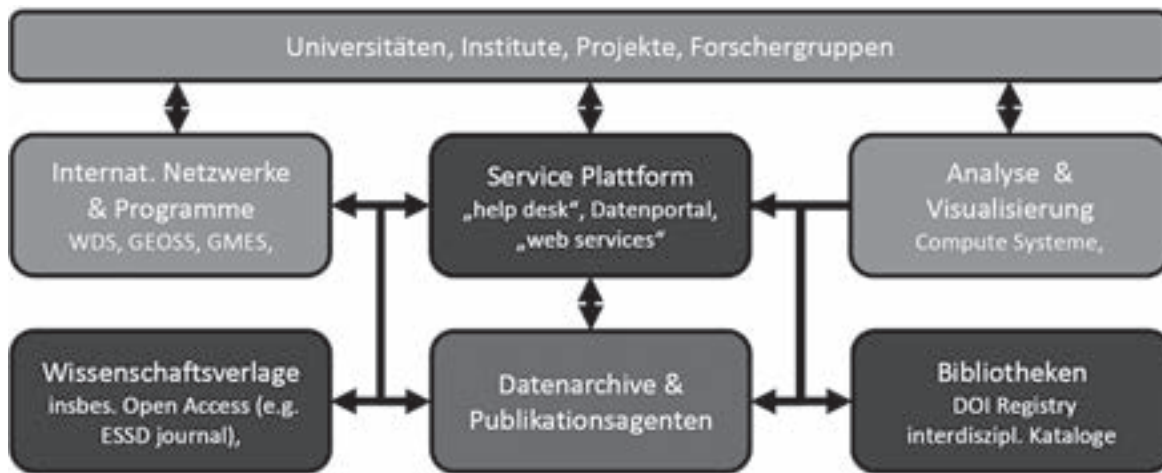


Abb.: Das Kompetenzzentrum als Kooperative von Datenarchiven, Wissenschaftsverlagen und Bibliotheken und der Service Plattform als Bindeglied zur Wissenschaft

schaffen, welche wissenschaftliche Projekte, Institute, Forschergruppen oder auch einzelne Wissenschaftler in allen Fragen des Datenmanagements begleitet – von der Projektplanungsphase über Datenerhebung, Qualitätssicherung, Registrierung und Langzeitarchivierung bis zur Publikation von Daten. Als technisches Kernstück wird eine entsprechende webbasierte personalisierbare Serviceplattform aufgebaut.

Ein wesentlicher Teil des Vorhabens konzentriert sich auf die Erarbeitung technischer und organisatorischer Zusammenhänge zwischen den beteiligten Datenarchiven, den assoziierten Verlagen und der TIB. Kritisch sind dabei einerseits die Abstimmung und technische Realisierung der redaktionellen Schnittstellen zwischen Verlagen und Datenarchiven und andererseits die Bereitstellung leistungsfähiger und zuverlässiger Web-Services für die Cross-Referenzierung von Daten und Publikationen sowie die Einbindung in übergeordnete globale Netzwerke.

Über die Serviceplattform und die gemeinsame Infrastruktur werden zudem standardkonforme Datenzugriffssysteme und Schnittstellen zur Verfügung gestellt. Dazu gehören ein Portal für die Datenrecherche und Dienste, die den Austausch von Metadaten ermöglichen (Mehrwertdienste zur allgemeinen Verfügung). Letzteres erlaubt z. B. via OAI-PMH<sup>22</sup> die Einbindung des Datenkataloges in die Europeana<sup>23</sup> oder in das »ISI Web of Knowledge«<sup>24</sup>. Letzteres ist Gegenstand einer laufenden Zusammenarbeit mit Thomson Reuters. Ziel ist die Schaffung eines Index für Datenpublikationen.

Das Vorhaben ist gut auf die Planungen für das ICSU WDS abgestimmt, nicht zuletzt wegen der aktiven Beteiligung der nationalen »World Data Center« am Aufbau des WDS (PANGAEA ist im WDS-Lenkungsausschuss vertreten). In einer späteren Phase sollen

daher Teile der Organisation und Infrastruktur in das WDS eingebunden werden.

Das WDS wie auch das beschriebene DFG-Vorhaben weisen konzeptionell in die richtige Richtung. Dennoch bleibt festzuhalten, dass gegenwärtig nur ein kleiner Teil der weltweit produzierten bio- und geowissenschaftlichen Daten wirklich nachnutzbar ist.<sup>25</sup> Daran haben auch die mittlerweile zahlreichen politischen Vorgaben und Richtlinien – z. B. durch die OECD<sup>26</sup>, die DFG<sup>27</sup> oder die Allianz der deutschen Wissenschaftsorganisationen<sup>28</sup> – wenig geändert. Zu den besonderen Herausforderungen der nächsten Jahre gehören:

Die nachhaltige Veränderung der Wissenschaftskultur im Umgang mit Daten. Dazu gehören nicht nur restriktive Vorgaben durch Fördereinrichtungen, sondern auch eine entsprechende Ausbildung des wissenschaftlichen Nachwuchses sowie die Schaffung von Anreizsystemen wie die Zitierfähigkeit von Datenpublikationen.

Datenintensive Wissenschaft erfordert nicht nur die allgemeine Verfügbarkeit von Daten, sondern darüber hinaus die Möglichkeit, Daten aus unterschiedlichsten Quellen effizient zu integrieren.<sup>29</sup> Dies wiederum erfordert einen hohen Grad an Interoperabilität datenbereitstellender Systeme und die Verwendung einheitlicher Standards für die Beschreibung und Strukturierung von Daten.

Nicht zuletzt erfordern Aufbau und Unterhalt von Informationsinfrastrukturen für die Wissenschaft erhebliche finanzielle Mittel. Diese sollten fester Bestandteil der Forschungsförderung sein, d. h. überall dort, wo Daten produziert werden, müssen auch Mittel für die Qualitätssicherung, Publikation und Archivierung mit eingeplant werden. Für die Publikation von Daten wäre die Finanzierung nach dem Open Access-Modell sinnvoll (»datacharges«).

**Zitierfähigkeit von Datenpublikationen**

**Forschungsdaten sichern, publizieren, archivieren**

## ANHANG

### **PANGAEA® – Data Publisher for Earth & Environmental Science (vormals WDC-MARE), MARUM Universität Bremen / AWI Bremerhaven, [www.pangaea.de](http://www.pangaea.de)**

Die breitspektral über die gesamten Geo-, Bio- und Umweltwissenschaften verteilten Datenbestände von PANGAEA dienen der Erforschung globaler Umweltveränderungen. Der Schwerpunkt liegt auf georeferenzierbaren Daten aus den Bereichen der Ozeanographie, Meeresgeologie, Paläoozeanographie und Meeresbiologie. Operationelle Plattform ist das Informationssystem PANGAEA. Das System beinhaltet gegenwärtig ca. 550.000 Datensätze mit mehr als fünf Milliarden Datenpunkten zu ca. 35.000 verschiedenen Parametern aus allen Weltmeeren und Kontinenten und finanziert sich mehrheitlich über Projektdatenmanagement und den Aufbau von Geodaten-Infrastrukturen (47 internationale, 51 EU – davon mehrere NoE, 42 nationale Projekte während der letzten zwölf Jahre, [www.pangaea.de/projects](http://www.pangaea.de/projects)). 2008 erhielt PANGAEA den »Computerworld Honest Award«. Ende 2008 wurde im Rahmen einer Kooperation mit Elsevier ein gemeinsamer Web-Service aufgebaut, welcher die direkte Referenzierung von PANGAEA-Datensätzen als Supplement zu wissenschaftlichen Publikationen über den Verlagskatalog (Science Direct) ermöglicht. Ähnliche Kooperationen wurden seitdem mit Springer, der American Geophysical Union (AGU), Copernicus, der Oxford University Press und Wiley initiiert. Seit 2010 harvestet Thomson Reuters PANGAEA Metadaten in das Web of Knowledge.

### **World Data Center for Remote Sensing of the Atmosphere (WDC-RSAT), DLR Oberpfaffenhofen, <http://wdc.dlr.de>**

Der Fokus liegt auf Daten, höherwertigen Informationsprodukten und anwendungsorientierten Services (z. B. Luftqualität, Energie, Gesundheit, Umweltkonventionsmonitoring, Naturkatastrophen), hauptsächlich basierend auf satellitenbasierten Messungen der Atmosphäre (Spurengase, Aerosole, Strahlung, Wolken, Dynamik). Neben eigenen Repositorien, die vornehmlich auf europäischen Missionen (z. B. ENVISAT, ERS-2, MetOp, MSG) beruhen, wird der Zugang zu praktisch allen existierenden satellitenbasierten Informationen weltweit ermöglicht (»one-stop-shop«). WDC-RSAT nimmt neben seiner Funktion als ICSU-WDC auch Mandate wahr als UN-WMO-WDC und in Kooperation mit NASA als Portal für CEOS-Atmosphäre. WDC-RSAT dient als Plattform für atmosphärenbezogene

GMES-Services von ESA und EU, setzt hierbei Verfahren zur Verknüpfung von Daten mit numerischen Modellen (Datenassimilation) ein und stellt Werkzeuge zur Verschneidung verschiedener Datenebenen zur Verfügung. WDC-RSAT ist Bestandteil der GEO-Themenbereiche »Disasters«, »Health«, »Energy« und »Climate«. WDC-RSAT dient als Daten- und Kommunikationsplattform für das internationale Meßnetzwerk NDMC und fungiert als Zugang zu satellitenbasierten Daten für bodengebundene Observatorien im Rahmen des WMO-Global Atmosphere Watch Programms.

### **World Data Center on Climate (WDC-C), DKRZ, Hamburg, [www.mad.zmaw.de/wdccc/](http://www.mad.zmaw.de/wdccc/)**

Das WDC-C konzentriert sich auf die Zusammenführung von Daten der Klima- und Erdsystemforschung. Schwerpunkt bilden die Ergebnisse aus numerischen Modellrechnungen und zugehörige Beobachtungsdatensätze zum Antrieb der Modelle und zur Verifikation ihrer Ergebnisse. Im Rahmen des WDC-C werden für internationale Programme Datenzentren gepflegt. Im IPCC Data Distribution Center (DDC) werden ausgewählte Daten internationaler globaler Klimamodelle zusammengeführt und zentral zur Verfügung gestellt. Diese Daten bilden die Grundlage für die alle drei bis fünf Jahre erscheinenden Assessment Reports, die als Referenz für das internationale Verständnis des Klimasystems gelten. Das WDC-C hat zurzeit einen Umfang von 400 Terabyte Klimadaten, die in ca. 1.350 Experimenten mit mehr als 160.000 Zeitreihen einzelner Variablen abgelegt sind. Alle Daten stehen über das Internet fein granular zum transparenten Zugriff bereit. Ca. acht Milliarden einzelne Datenentitäten sind gezielt und direkt abrufbar.

### **Helmholtz-Zentrum Potsdam Deutsches GeoForschungszentrum (GFZ), Potsdam, [www.gfz-potsdam.de](http://www.gfz-potsdam.de)**

Das GFZ erforscht als nationales Forschungszentrum für Geowissenschaften weltweit das »System Erde« mit den geologischen, physikalischen, chemischen und biologischen Prozessen, die im Erdinneren und an der Oberfläche ablaufen. Durch die erfolgreiche Forschungstätigkeit des GFZ sind seit 1992 viele wertvolle, unikale Datenbestände gewonnen worden, die für die Zukunft der Geowissenschaften ein erhebliches Potenzial besitzen. Die Datenbestände des GFZ umfassen Daten aus geodätischen Satellitenmissionen, seismologischen Messnetzwerken, geomagnetischen Observatorien, wissenschaftlichen Bohrprojekten und vielen weiteren Datensätzen aus der geowissenschaft-

lichen Forschung. Viele dieser Daten sind bereits als Datenveröffentlichungen für die wissenschaftliche Community zugänglich und werden über DOI referenziert. Im Rahmen der AG »Vertrauenswürdige Archive« im Kompetenznetzwerk Langzeitarchivierung (nestor) war das GFZ Potsdam an der Entwicklung des »Kriterienkatalogs vertrauenswürdige digitale Langzeitarchive« beteiligt. Projektdatenmanagement wird als gemeinsamer Dienst des GFZ, der »Geoinformation Technology« (CeGIT) und der GFZ Bibliothek Albert Einstein angeboten. Dieser Dienst umfasst Datenpflege, -publikation und Langzeitarchivierung sowie die Datenbereitstellung über Projektportale und virtuelle Forschungsumgebungen.

### Technische Informationsbibliothek (TIB), Hannover, [www.tib-hannover.de](http://www.tib-hannover.de)

Die TIB ist die deutsche zentrale Fachbibliothek für alle Gebiete der Technik sowie Architektur, Chemie, In-

formatik, Mathematik und Physik. Seit 2005 ist die TIB nicht-kommerzielle DOI-Registrierungsagentur für Forschungsdaten und war damit weltweit die erste Einrichtung dieser Art. Derzeit registriert sie für 33 Datenzentren aus Deutschland, Europa und den USA Forschungsdaten, graue Literatur und weitere nicht-textuelle Objekte und sichert so die Verfügbarkeit und Referenzierbarkeit von wissenschaftlichen Daten (bislang mehr als 750.000 registrierte Datensätze). Unter Federführung der TIB ist im Dezember 2009 der international besetzte Verein DataCite gegründet worden. Ziel ist die Registrierung von Forschungsdaten auf der Basis nationaler Zuständigkeiten. DataCite hat zwölf Mitglieder aus neun Ländern (Stand: Februar 2010). Seit 2006 ist die TIB auch Mitglied des Executive Board der International DOI Foundation (IDF). Im Jahr 2010 übernahm DataCite diese Mitgliedschaft der TIB und stellt außerdem den Vorsitzenden des IDF-Board.

<sup>1</sup> Wissenschaftsrat (2000), Thesen zur künftigen Entwicklung des Wissenschaftssystems in Deutschland, [www.wissenschaftsrat.de/download/archiv/pm\\_1400.pdf](http://www.wissenschaftsrat.de/download/archiv/pm_1400.pdf)

<sup>2</sup> Wissenschaftsrat (2011), Übergreifende Empfehlungen zu Informationsinfrastrukturen, [www.wissenschaftsrat.de/download/archiv/10466-11.pdf](http://www.wissenschaftsrat.de/download/archiv/10466-11.pdf)

<sup>3</sup> [www.earthobservations.org/](http://www.earthobservations.org/)

<sup>4</sup> [www.wmo.int/pages/themes/wis/index\\_en.html](http://www.wmo.int/pages/themes/wis/index_en.html)

<sup>5</sup> <http://icsu-wds.org/>

<sup>6</sup> [http://ec.europa.eu/research/infrastructures/index\\_en.cfm?pg=esfri](http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri)

<sup>7</sup> [www.iode.org/](http://www.iode.org/)

<sup>8</sup> [www.gbif.org/](http://www.gbif.org/)

<sup>9</sup> [http://ec.europa.eu/gmes/index\\_en.htm](http://ec.europa.eu/gmes/index_en.htm)

<sup>10</sup> <http://inspire.jrc.ec.europa.eu/>

<sup>11</sup> [www.icsu-fags.org/](http://www.icsu-fags.org/)

<sup>12</sup> Reference Model for an Open Archival Information System (OAIS) – (corresponds to ISO 14721:2003), <http://public.ccsds.org/publications/archive/650x0b1.pdf>

<sup>13</sup> S. Dobratz et al. (2009), Catalogue of Criteria for Trusted Digital Repositories, nestor materials, Deutsche Nationalbibliothek, Frankfurt (Main), Germany, <http://nbn-resolving.de/urn:nbn:de:0008-2010030806>

<sup>14</sup> <http://www.ipcc.ch/>

<sup>15</sup> [www.iodp.org/](http://www.iodp.org/)

<sup>16</sup> [www.icdp-online.org/](http://www.icdp-online.org/)

<sup>17</sup> <http://datacite.org/>

<sup>18</sup> Elsevier and PANGAEA Take Next Step in Connecting Research Articles to Data, press release (2010), [www.elsevier.com/wps/find/authored\\_newsitem.cws\\_home/companynews05\\_01616](http://www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_01616)

<sup>19</sup> W3C – World Wide Web Consortium [www.w3.org/](http://www.w3.org/)

<sup>20</sup> ISO – International Organisation for Standardization [www.iso.org/](http://www.iso.org/)

<sup>21</sup> OGC – Open Geospatial Consortium [www.opengeospatial.org/](http://www.opengeospatial.org/)

<sup>22</sup> [www.openarchives.org/pmh/](http://www.openarchives.org/pmh/)

<sup>23</sup> [www.europeana.eu](http://www.europeana.eu)

<sup>24</sup> <http://isiwebofknowledge.com/>

<sup>25</sup> Bryn Nelson (2009), Empty archives, Nature, Vol. 461, S. 160–163; Eefke Smit (2011), Abelard and Héloïse: Why Data and Publications Belong Together, D-Lib Magazine, Vol. 17, No. 1/2, [www.dlib.org/dlib/january11/smit/01smit.html](http://www.dlib.org/dlib/january11/smit/01smit.html)

<sup>26</sup> OECD (2007), OECD Principles and Guidelines for Access to Research Data from Public Funding, [www.oecd.org/dataoecd/9/61/38500813.pdf](http://www.oecd.org/dataoecd/9/61/38500813.pdf)

<sup>27</sup> DFG-Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (2009), Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten, [www.dfg.de/forschungsfoerderung/wissenschaftliche\\_infrastruktur/lis/veroeffentlichungen/dokumentationen/download/ua\\_inf\\_empfehlungen\\_200901.pdf](http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/veroeffentlichungen/dokumentationen/download/ua_inf_empfehlungen_200901.pdf)

<sup>28</sup> Allianz der deutschen Wissenschaftsorganisationen (2010), Grundsätze zum Umgang mit Forschungsdaten, [www.allianzinitiative.de/fileadmin/user\\_upload/Home/Video/Grundsätze%20Umgang%20mit%20Forschungsdaten.pdf](http://www.allianzinitiative.de/fileadmin/user_upload/Home/Video/Grundsätze%20Umgang%20mit%20Forschungsdaten.pdf)

<sup>29</sup> Jim Gray on eScience: a transformed scientific method, in: Tony Hey, Stewart Tansley, Kristin Tolle (Eds.), The Fourth Paradigm: Data-Intensive Scientific Discovery [http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th\\_paradigm\\_book\\_complete\\_lr.pdf](http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf)

## DIE VERFASSER

**Michael Diepenbroek** ist wissenschaftlicher Mitarbeiter des MARUM an der Universität Bremen und Koordinator von PANGAEA®, MARUM – Zentrum für Marine Umweltwissenschaften der Universität Bremen, Leobener Straße, 28359 Bremen, Mail: [mdiepenbroek@pangaea.de](mailto:mdiepenbroek@pangaea.de)

**Prof. Dr. Dr. h.c. Gerold Wefer** ist Professor für Allgemeine Geologie an der Universität Bremen und Direktor des DFG-Forschungszentrums »Der Ozean im System Erde«-MARUM, MARUM – Zentrum für Marine Umweltwissenschaften der Universität Bremen, Leobener Straße, 28359 Bremen, Mail: [gwefer@marum.de](mailto:gwefer@marum.de)