



se nicht nur für Materialwissenschaftler oder Architekten interessant, sondern zum Beispiel auch für Erdbebenforscher (Originalveröffentlichung: Markus J. Buehler, Huajian Gao: Dyna-

mical fracture instabilities due to local hyperelasticity at crack tips. *Nature*, 19 January 2006).

Monika Huch, Adelheidsdorf

Der Wert von Daten liegt in ihrer Nutzung

Datenbanken haben einen schlechten Ruf. Erst hat man, erzwungen durch obskure Formate, ein Problem Daten hineinzubekommen, dann wissen die anderen nicht, wie man sie wieder findet, geschweige denn herausbekommt. Findet man Daten, schwächelt die Dokumentation, wenn denn überhaupt vorhanden. Auch weigern sich viele Wissenschaftler, gebrannt durch eine Konfrontation mit inakzeptablen Qualitätsproblemen, überhaupt mit recycelten Zahlen weiterzuarbeiten. Und dann gibt es da noch die so genannten Portale und Metadatenbanken: In dem Glauben genau das zu finden, was man braucht, klickt man sich die Finger wund, um schließlich auf der *Homepage* eines Institutes zu laden oder die Irrfahrt durch das Internet mit der Meldung *Error 404, File not found* ein abruptes Ende findet. Sollte durch eine Verkettung glücklicher Umstände die Datennahrungskette doch einmal bis zu einer Publikation geklappt haben (der Grund sind meist die menschlichen und nicht die elektronischen Netzwerke), finden sich in der Literaturliste die Datenautoren bestenfalls fragmentarisch. Diese können sich mit einer *frdl. mdl. Mittl.* oder einer Erwähnung in den *Acknowledgements* begnügen. Mal ehrlich: Wie soll unter diesen Voraussetzungen eine wissenschaftlich sinnvolle und eigentlich dringend notwendige Datenarchivierung den ewigen Kinderschuhen und Spielkreisen entwachsen?

Einige Leser werden jetzt erwidern, dass es doch das System xyz gibt, dass man seinerzeit da und dort Datenbanken aufgebaut habe, oder man argumentiert, dass in den USA entsprechende Systeme zur Verfügung stünden. Sicherlich - viele Bits sind seit dem ersten Loch in einer Pappkarte erfasst worden, viele Systeme wurden technisch initiiert. Allein der Begriff ‚Ar-

chivierung‘ beinhaltet zwingend das Adjektiv ‚langfristig‘ und davon ist man mit der Einführung von *Datenmanagement auf Webservern* weiter entfernt, denn je. Der Enthusiasmus beim Aufbau entsprechender Systeme endet beim Übergang in den Dauerbetrieb – totale Ernüchterung beim Versuch an Daten zu kommen – der Initiator wird von einem neuen Arbeitgeber oder dem Ruhestand gerufen. Den Rest erledigen technische Sachzwänge, wenn die nächste Generation an Speichermedien, Hardware und Software einen nicht mehr zu bewältigenden Aufwand fordern oder die Technik des Backup schlichtweg ignoriert wurde.

Aus den oben dargestellten Hinderungsgründen ergeben sich technische und organisatorische Konsequenzen wenn denn der Datenfluss in entsprechende System verbessert werden soll – in Systeme, die im bibliothekarischen Sinne einer nachhaltigen Verfügbarkeit, eines Mehrwertes und der Nachnutzbarkeit von wissenschaftlichen Inhalten gerecht werden. (1) Die Akzeptanz eines Datensystems steht und fällt mit der einfachen Findbarkeit, dem Zugriff und dem Verbreitungsgrad seiner Inhalte. (2) Die Daten müssen mit standardisierten Beschreibungen versehen sein, damit der Nutzer ihre Qualität und Herkunft bewerten kann (keine Daten ohne Metadaten, keine Metadaten ohne Daten). (3) Wissenschaftler wären motivierter Daten zu liefern, wenn Sie dafür entsprechend zitiert würden. Jeder Datensatz muss somit in der Beschreibung ein bibliographisch verwendbares Zitat beinhalten. Teil des Zitates sollte, wie heute bei etablierten Verlagen üblich, ein dauerhafter Identifikator (z.B. DOI) sein. (4) Um die Nachhaltigkeit zu gewährleisten, muss Datenhaltung durch etablierte Zentren und Systeme erfolgen, die die ‚technischen Sachzwänge‘ im Griff ha-



ben. (5) Neben dem Datenzitat gewinnt die Datenpublikation, auch in Verbindung mit einem *peer-review*, in der internationalen Diskussion zunehmend an Bedeutung. (6) Zuwendungsgeber, Institute und Projekte sollten in entsprechenden Erklärungen und Bestimmungen ihre Datenpolitik formulieren.

Motivation, sich diesen Herausforderungen zu stellen, ist ausreichend vorhanden. Die DFG fordert zur Sicherung guter wissenschaftlicher Praxis eine sachgerechte Archivierung der den Publikationen zugrunde liegenden Primärdaten. Der offene Zugang (*Open Access*), in der Budapester Initiative begründet, wurde mit der Berliner Erklärung für wissenschaftliche Ergebnisse im weitesten Sinn und damit auch für Daten formuliert. Unterstützung erfährt die moralische Verpflichtung zur Datenarchivierung einmal mehr, wenn die OECD demnächst ihre ‚Empfehlungen zum Zugang zu Forschungsdaten, die mit Mitteln der öffentlichen Hand erhoben wurden‘ herausgeben wird.

Moral ist die eine Seite. Die Motivation einen kleinen Beitrag zu leisten für eine bessere und effizientere Wissenschaft, eventuell sogar helfen neue Fragen zu lösen, hat einen weiteren signifikanten Stellenwert. Gerade Systeme mit georeferenzierten Daten, wenn denn wirklich in konsistenter Form archiviert, ermöglichen eine einfache Suche mit Überblick und Zugriff auf das was existiert - in Zeiten exponentiellen Wachstums von Daten und Publikationen unabdingbar. Zum anderen erhalten entsprechende Datensammlungen einen wissenschaftlichen Mehrwert, mit dem sich neue umfassendere Fragestellungen angehen lassen.

In Deutschland bemüht sich eine Gruppe von Geowissenschaftlern und Physikern in einer fruchtbaren Kooperation zwischen einem Helm-

holtz-Forschungszentrum (AWI) und der Universität Bremen (MARUM) den Anforderungen einer sachgerechten Haltung georeferenzierter Daten der Erdsystemforschung gerecht zu werden. Das aus einem kleinen Projekt in 12 Jahren entstandene Informationssystem PANGAEA bietet heute vieles, was zurzeit technisch möglich ist. Die Verknüpfung aller Anforderungen an Langzeitarchiv, Datenbibliothek und Publikationssystem mit einem flexiblen, generischen Datenmodell im Hintergrund ist wohl einzigartig im Internet. PANGAEA konzentriert sich auf Inhalte und stellt somit auch einen ersten Vorrat an Daten zur Verfügung. Dieser scheint im internationalen Vergleich mit anderen Systemen zwar beträchtlich, in Relation zu dem was täglich produziert wird, mahnt er allerdings zur Bescheidenheit.

Auch wenn man von einer durchgängigen Akzeptanz und der konsequenten Nutzung durch die Wissenschaft noch kurze geologische Zeiträume entfernt ist: Mit dem Aufbau und Betrieb einer ‚Datenbibliothek‘ für die Erdsystemforschung stellen sich die Betreiber den Herausforderungen und Aufgaben einer modernen Erfassung und Verbreitung geowissenschaftlicher Daten. Wirklich langfristig Erfolg haben werden solche Systeme allerdings nur mit der Unterstützung aller an der Forschung beteiligten Partner: Zuwendungsgeber, Projekte, Wissenschaftsorganisationen, Fachgesellschaften, Institute und jedes einzelnen Wissenschaftlers. Der Wert von Daten liegt in ihrer Nutzung (*Zitat ICSU/CODATA*); eine Nutzung ist allerdings nur möglich, wenn die Daten auch verfügbar sind.
info@pangaea.de

Hannes Grobe, Bremerhaven & Michael Diepenbroek, Bremen

„Balkanisierung“ der Geo-Studiengänge? – Diskussionen in EFG und Geokonferenz

Die deutschen Universitäten sind im „Bologna-Fieber“. Der politisch forcierte Bologna-Prozess

führt dazu, dass in Deutschland und Europa weitflächig Studiengänge mit Bachelor-Master-