

OFFENER ZUGANG ZU DATEN

# Quantensprung in der Kooperation

Daten und andere Quellenmaterialien sollten als Teilergebnis des Forschungsprozesses offen zugänglich gemacht werden

**Ob es sich um Ergebnisse klinischer Versuche, die Klimaforschung, sozialwissenschaftliche Statistiken oder den Datenstrom einer Sonnensonde handelt: Data Sharing wird zukünftig in vielen Disziplinen zum Alltag gehören und den Mehrwert internationaler Kooperationen potenzieren. Bis dahin sind allerdings noch eine Reihe technischer und organisatorischer Fragen zu lösen.**

Jedes Jahr werden große Summen in die Gewinnung wissenschaftlicher Daten investiert. Dennoch werden viele dieser Daten nur partiell genutzt, da sie in der Regel lediglich einer begrenzten Gruppe von Forschern zugänglich sind. Der Forschungsprozess würde durch Data Sharing sowohl beschleunigt als auch qualitativ verbessert, da viele Ergebnisse nur anhand der verwendeten Daten nachvollzogen und für Fol-

geprojekte verwendet werden können. Neben der Förderung von Kooperationen ist die rechtzeitige Aufdeckung von Fälschungen ein weiteres Argument für die Offenlegung von Daten. Aus diesem Grund empfiehlt die Deutsche Forschungsgemeinschaft bereits seit 1998 in ihren Richtlinien für die gute wissenschaftliche Praxis die Aufbewahrung von Primärdaten für zehn Jahre. Eine Reihe anderer Institutionen hat sich dieser Forderung inzwischen angeschlossen. Auch die von den großen deutschen Forschungsorganisationen unterzeichnete Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen von 2003 bezieht in ihre Definition von Open Access ausdrücklich Rohmaterialien, Metadaten und Quellenmaterial mit ein. Ein Weg zur Umsetzung von Data Sharing könnte darin bestehen, dass von Seiten der Mittelgeber eine entsprechende Verpflichtung

ausgesprochen wird. Vor diesem Hintergrund bereitet beispielsweise die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) gegenwärtig eine Empfehlung an ihre Mitglieder vor, den offenen Zugang zu Rohdaten aus öffentlich geförderter Forschung gesetzlich zu verankern.

Einige Disziplinen machen bereits seit Jahren positive Erfahrungen auf diesem Gebiet. So ist es unter anderem in der Astrophysik üblich, Zeitschriftenartikel mit einem Link zu versehen, unter dem das verwendete Datenmaterial einzusehen ist. Auch Publikationen über Gensequenzen werden nicht akzeptiert, bevor die Sequenz nicht in einer anerkannten Datenbank hinterlegt wurde.

Dennoch gibt es noch eine Reihe von Problemen, die sowohl technische und organisatorische Aspekte betreffen als auch die Wissenschaftskultur selber. Zu den noch ungelösten Fragen gehört zum Beispiel die Sicherstellung von Findbarkeit, Vertrauenswürdigkeit und Persistenz der Datenquellen. Ortsangaben im Internet sind in den meisten Fällen nach kurzer Zeit ungültig. Auch die Finanzierung für die Aufbereitung und Einstellung der Daten, die eventuell notwendige Software und technische Wartung ist nicht geklärt.

Hinderlich für die Motivation der Forscher wiederum kann sich hingegen die Praxis des „publish or perish“ auswirken. Zurzeit wird in den meisten Disziplinen die Publikation von Daten weder verlangt noch honoriert. Im Gegenteil – wer der sorgfältigen Aufarbeitung und Veröffentlichung seiner Daten Zeit widmet, statt sich dem nächsten Zeitschriftenbeitrag zuzuwenden, ist im Vergleich zu Kollegen benachteiligt. Auch besteht oftmals die Befürchtung, dass interessante Daten, in deren Gewinnung und Verifikation viel Arbeit investiert wurde, nach



*Ein Projekt in den Geisteswissenschaften, dessen Daten frei zugänglich sind, ist das Pantheon-Projekt des Karman Centers der Universität Bern. Archäologen, Ingenieure und Historiker haben das Pantheon in Rom mit modernster Technologie vermessen und auf dieser Grundlage eine komplette, dreidimensionale Darstellung des Bauwerks in digitaler Form erstellt.*

der Veröffentlichung von anderen Wissenschaftlern ohne Quellenverweis genutzt werden könnten. Deswegen entscheiden sich Forscher immer wieder doch dafür, ihre Daten für sich zu behalten und sie selbst zu verwerten. So lange hier nicht eine Umorientierung erfolgt und die Bereitstellung nachnutzbarer Daten nicht grundsätzlich als eigenständige, zitierbare und bewertete wissenschaftliche Leistung Anerkennung findet, wird sich Data Sharing nicht flächendeckend durchsetzen.

### Rechtliche Fragen

Um solchen Vorbehalten zu begegnen, werden inzwischen in zahlreichen Projekten Maßnahmen getroffen, die den Urhebern der Daten exklusive Nutzungsrechte für einen gewissen Zeitraum einräumen. In der Praxis erweisen sich diese Regelungen jedoch oftmals als unzureichend, da es keinen rechtsverbindlichen Rahmen gibt. Ein deutlicher Fortschritt ist die Empfehlung, das Lizenzsystem der Creative Commons auch für die Wissenschaft zu übernehmen. Bei Creative Commons handelt es sich um eine weltweit agierende Non-Profit-Organisation, die Autoren und Künstler dabei unterstützt, ihre Werke im Internet zu verbreiten. Zu diesem Zweck veröffentlicht sie Standard-Lizenzverträge, mit denen die Urheber ihre Werke schützen können. In einzelnen Bereichen ist das Modell der Creative Commons nicht ausreichend und wurde durch die Science Commons ergänzt. Diese Organisation fördert Open Access, indem sie rechtliche Probleme der digitalen Dissemination identifiziert und Lösungsansätze anbietet.

In manchen Fachgebieten wie beispielsweise in der Medizin sind neben dem Urheberrecht Fragen des Datenschutzes zu berücksichtigen. Hier sollten gesetzliche Regelungen geschaffen werden, die die Interessen des einzelnen Patienten schützen, und die Bereitstellung von anonymisierten Rohdaten ermöglichen. In anderen Bereichen, wie bei den Geobasisdaten der staatlichen Vermessungsämter, den Satellitenbildern des Deutschen Fernerkundungsdatenzentrums oder den meteorologischen Daten des Deutschen Wetterdienstes wird Data Sharing durch das gegenwärtig bestehende Finanzierungsmodell verhindert. Hier gibt es Vorga-



*Forscher mit Eisbärabwehr vor dem Forschungsschiff „Polarstern“ in der Arktis. Die Primärdaten aus den Projekten des Internationalen Polarjahrs von 2007 bis 2009 werden offen zugänglich gemacht werden. Sogar die Rohdaten sollen bereits während der Expedition zur Koordination der über 250 Projekte zur Verfügung stehen.*

ben, wonach ein wesentlicher Teil der Kosten aus dem Verkauf der Daten zu decken ist. Alternative Geschäftsmodelle fehlen. Es wäre durchaus zu bedenken, ob der gesamtwirtschaftliche Nutzen nicht größer wäre, wenn die Daten kostenlos oder zu einem nominalen Preis bereitgestellt würden. So könnten sich hier ganz neue Dienstleistungen entwickeln, zum Beispiel spezialisierte Wettervorhersagen.

### Internationales Polarjahr

Ein großes Projekt zum Data Sharing beginnt mit dem Internationalen Polarjahr, das von März 2007 bis März 2009 dauert. Es schließt jeweils zwei arktische und zwei antarktische Sommer ein und ermöglicht so an beiden Polen zwei aufeinander folgende Messkampagnen. Umfangreiche internationale Forschungsprojekte widmen sich in diesem Zeitraum der intensiven Erforschung der Polargebiete und ihrer Bedeutung für das Klimageschehen. Mit 50.000 Forschern aus über 60 Ländern handelt es sich um das größte Kooperationsprojekt, das es je gab. Die gewonnenen Daten sollen so früh wie möglich offen zugänglich gemacht werden, auch dann, wenn sie noch nicht endgültig ausgewertet sind. Die Ergebnisse dieses Programms werden eine Art Schnappschuss des Planeten

liefern, der der Forschung kommender Jahrzehnte als Grundlage und Referenzsystem dienen soll.

Wie die Urheberrechts-, Datenschutz- und Finanzmodelle zum Data Sharing auch im Einzelnen aussehen werden: Der offene Zugang zu Primärdaten ist sinnvoll, wird einen erheblichen Mehrwert erbringen und aus der Forschung der Zukunft nicht mehr wegzudenken sein.

**Dr. Hans Pfeiffenberger**  
Alfred-Wegener-Institut für Polar- und Meeresforschung, Bremerhaven

**Dr. Jens Klump**  
GeoForschungsZentrum Potsdam