

Open Access – nur Texte oder auch Primärdaten?

Jochen Fahrenberg

revidierte Fassung 5 .10. 2009

Plädoyers für Open Access

Die Aufrufe und Aktivitäten zum Thema Open Access sind zu einer breiten Strömung geworden: On-line Journals, Forschungsberichte, Dissertationen und – noch sehr selten – komplette *e-Bücher*. Bemerkenswert ist jedoch, dass es sich fast ausschließlich um Texte handelt, höchstens am Rande werden die primären Forschungsdaten erwähnt. Dass bereits vor fast 40 Jahren ein von der *American Psychological Association* herausgegebener *Journal Supplement Abstract Service* die Archivierung von sonst nicht leicht zu publizierenden Texten und Materialien ermöglichte und anfänglich auch an Datensätze gedacht war, scheint weitgehend vergessen zu sein. Der JSAS kann als einer der Vorläufer computer-gestützter Dokumentationssysteme angesehen werden. Heute werden in den USA Zeitschriften-Artikel und andere Publikationen zunehmend durch „Supplementary Online Material“ ergänzt, doch handelt es sich dabei in der Regel noch nicht um die Primärdaten. (1) *Sharing Research Data* war bereits 1985 Thema eines Buches von Fienberg, Martin und Straf. – Heute stehen die Archivierung und die „Nachnutzung“ der Forschungsdaten wieder auf der Tagesordnung.

Open Access stützt sich auf die Kommunikationsmöglichkeiten des Internet. Dementsprechend sind auch zahlreiche Links zu Grundsatzserklärungen und verschiedenen Initiativen zu finden sowie Links zu einer wachsenden Anzahl von Datenbanken (Open-Access Data Repositories).

Wer sich dagegen einige typische Homepages von Instituten der Psychologie und die Seiten der Mitarbeiter ansieht, wird die Publikationslisten, eventuell einige preprints und postprints oder noch Forschungsberichte finden, jedoch keine Hinweise auf die zugrundeliegenden Datensätze. Was wären all diese Texte ohne ihre empirische Basis, ohne die primären Daten? Was spricht in der Wissenschaft gegen das allgemeine Prinzip von „data-enriched“ Zeitschriftenartikeln, bei deren Publikation die wichtigen Datensätze u.a. Informationen auf einem Datenserver archiviert werden?

Nachdem sich Forschungsorganisationen für die Prinzipien des Open Access von wissenschaftlichen Texten eingesetzt haben, dehnt sich dieser Appell auch auf den Zugang zu *Primärdaten* aus. Im Fach Psychologie nimmt die American Psychological Association APA zwar eine progressive Haltung ein, doch mangelt es noch an der praktischen Umsetzung, d.h. an Regeln und geeigneten Data Repositories. Dabei fällt auf, dass sich die Diskussion noch kaum mit den psychologischen Tests und der Dokumentation spezieller Methoden befasst hat. Die Archivierung der Konstruktions- und Normierungsdaten psychologischer Tests blieb weitgehend ausgeklammert. Die Vermutung, es könne an der kommerziellen Seite der Testpublikation liegen, bleibt noch zu diskutieren. Schließen Publikation und Verkauf eines Tests aus, dass die primären Testdaten zugänglich gemacht werden?

Hier werden neuere Diskussionsbeiträge und wichtige Links zusammengefasst, um einen Eindruck zu vermitteln, wie es, insbesondere in der Psychologie, um den freien Zugang zu Primärdaten der Forschung („open data“, „data sharing“) steht.

Wichtige Initiativen und Richtlinien

Die Entwicklung der Idee des Open Access in den Wissenschaften hat sich in einer Reihe von Appellen, Denkschriften und Initiativen manifestiert:

[Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities](#)

[Budapest Open Access Initiative](#)

Deutsche Forschungsgemeinschaft: [Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten von Wissenschaftlern unter besonderer Berücksichtigung von Open Access](#)

[Deutsche Initiative für Netzwerkinformationen e.V.](#)

[Deutsche UNESCO-Kommission: Open Access. Chancen und Herausforderungen](#)

[OECD Committee for Scientific and Technological Policy](#)

Diese Beiträge und Resolutionen sind meist auf sehr hohem Allgemeinheitsniveau formuliert und beziehen sich nahezu ausschließlich auf die Publikation von *Texten*. (Das an der Universität von Lund in Schweden verwaltete [Directory of open access journals](#) verzeichnet gegenwärtig 4361 Zeitschriften.)

Initiativen zur Archivierung von primären Daten

Die amerikanische *National Science Foundation, Division of Social and Economic Sciences*, publizierte Richtlinien zur [Data Archiving Policy](#), ebenso der *Council on Governmental Relations: Access to and retention of research data: Rights and responsibilities (2006). Hier werden auch die Probleme, die Motive und die Hindernisse der Archivierung kommentiert. Bewilligungsempfänger aller Fachbereiche sollen spezielle Pläne entwickeln und vorlegen, wie die mit Unterstützung der NSF erhobenen Daten anderen Wissenschaftlern zugänglich gemacht werden.*

„The purpose of this policy is to advance science by encouraging data sharing among researchers. Data sharing strengthens our collective capacity to meet scientific standards of openness by providing opportunities for further analysis, replication, verification and refinement of research findings. These opportunities enhance the development of fields of research and support the potential for cross-directorate activity. In addition, the greater availability of research data will contribute to improved training for graduate and undergraduate students, and make possible significant economies of scale through the secondary analysis of extant data. Finally, researchers have a special obligation to scientific openness and accountability when the research is publicly funded.

SES supports a wide range of disciplines. Therefore, the nature of the data, the way they are collected, analyzed, and stored, and the pace at which this reasonably occurs vary widely. There are different storage facilities and different access requirements for, e.g., large-scale survey data, oral interviews with scientists and other subjects, and data generated by experimental research. **Grantees from all fields will develop and submit specific plans to share materials collected with NSF support, except where this is inappropriate or impossible.** These plans should cover how and where these materials will be stored at reasonable cost, and how access will be provided to other researchers, generally at their cost.” (NSF [Data Archiving Policy](#))

Im anschließenden Text werden unterschiedliche Datenkategorien angesprochen: Quantitative Social and Economic Data Sets, Qualitative Information, Experimental Research, Mathematical and Computer Models, ohne jedoch ins Detail der notwendigen Dokumentationssysteme zu gehen. Anstelle einer zentralen Forschungsdatenbank scheint eine regionale bzw. fachliche Aufgliederung angestrebt zu werden. In vielen Bereichen der Naturwissenschaften und Technik ist diese Datenarchivierung bereits in sehr viel breiterer Weise aufgebaut worden, z.B. in Astronomie, Biologie, Geowissenschaften und Physik.

Beobachtern fielen beim Vergleich naturwissenschaftlicher Fächer unterschiedliche Einstellungen und Gewohnheiten im Umgang mit Primärdaten auf. In diesen Disziplinen sind viele Arbeitsrichtungen kaum noch ohne ein übergreifendes und auch internationales data sharing möglich. In anderen Bereichen zeigte sich, dass die Gründung eines Datenarchivs allein noch nicht ausreicht; die Archive müssen zweckmäßig gefüllt, organisiert und akzeptiert werden (Nelson, 2009).

Die Deutsche Forschungsgemeinschaft veröffentlichte im Januar 2009 die vom *Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme* verfassten [Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten](#).

„Die nun vorliegenden Empfehlungen richten sich sowohl an die Wissenschaftler als auch die Informationsspezialisten. Es soll damit das Bewusstsein für die Dringlichkeit dieses Problems geweckt werden. Darüber hinaus werden die Ziele benannt, die bei der Umsetzung berücksichtigt werden sollten und deren Einhaltung von grundlegender Bedeutung ist. Schließlich sollen die Wissenschaftler dazu angeregt werden, dieser Aufgabe auf dem Wege der Selbstverpflichtung nachzukommen und für eine sachgemäße Umsetzung im Rahmen der gegebenen Strukturen in eigener Initiative Sorge zu tragen.“ „Auf der Grundlage dieser Empfehlungen wird die DFG im Rahmen ihrer Fördermöglichkeiten Mittel bereitstellen, um den Ausbau der Speicher- und Nachweisstrukturen zu unterstützen. Insbesondere regt sie zu Explorations-Projekten an, um für die einzelnen Disziplinen weitere Erfahrungen zu sammeln. Die Regelung der Aufbewahrung und Bereitstellung der Daten sollte von Anfang an im Arbeitsplan des Wissenschaftlers vorgesehen und im Rahmen der Förderverfahren bereits in den Anträgen dargelegt werden.“

Zwei der sieben Empfehlungen werden hier zitiert:

2. Es ist ein fachspezifisches Organisationskonzept festzulegen, das die nachhaltige Speicherung der Daten regelt. Die Wissenschaftler bestimmen jeweils innerhalb ihrer Disziplinen eine Institution oder organisatorische Struktur, die sowohl die rasche, komfortable Bereitstellung der Daten im laufenden Forschungsprozess als auch die langfristige Verfügbarkeit sicherstellt. Darüber hinaus ist zu vereinbaren, ob es eine maximale Aufbewahrungsfrist geben soll, nach deren Ablauf die Daten dem Forscher zurückgegeben oder gelöscht werden.

5. Jeder Wissenschaftler stellt seine Forschungsprimärdaten nach Möglichkeit überregional und frei zur Verfügung. In Abhängigkeit von der jeweiligen Fachkultur soll unter den Wissenschaftlern Konsens darüber hergestellt werden, dass die Daten unmittelbar nach Abschluss der Forschungen oder nach wenigen Monaten der Öffentlichkeit frei zur Verfügung gestellt werden. Sofern die Daten in Projekten erarbeitet wurden, die aus öffentlich-rechtlichen Mitteln finanziert wurden, stehen sie im Grundsatz der Öffentlichkeit frei zur Verfügung. Für Forschungsdaten, die der Patentierung oder anderer unmittelbarer wirtschaftlicher Nutzung unterliegen, sind gesonderte Regelungen zu treffen. Dieses soll im Konsens mit den betroffenen Wissenschaftlern geschehen.“

Von den deutschen Wissenschaftsorganisationen sind weitere *Grundsätze* zum Umgang mit Forschungsdaten zu erwarten. Die *speziellen Regelungen* wären jedoch von den einzelnen Fächern auszuarbeiten.

Nicht auf die sekundäre Nutzung von Daten, sondern primär auf die damit ermöglichte Kontrolle der wissenschaftlichen Zuverlässigkeit bezieht sich die Empfehlung der Deutschen Forschungsgemeinschaft, dass die Institute (nicht die Autoren) die Primärdaten, die Grundlage von Veröffentlichungen waren, für 10 Jahre „auf haltbaren und gesicherten Datenträgern“ aufbewahren (Empfehlung 7 der *Vorschläge zur Sicherung guter wissenschaftlicher Praxis*, DFG, 1998).

Weitere Einblicke in die gegenwärtige deutsche Diskussion vermitteln der Beitrag von Pfeiffenberger (2008) [Offener Zugang zu wissenschaftlichen Primärdaten](#) und der von Pampel (2008) verfasste [Bericht über den Workshop Offener Zugang zu Daten – eine Herausforderung im Rahmen der Open-Access-Tage 2008 in Berlin](#). Eine umfangreiche Sammlung von Links zu [Data Sharing and Replication](#) geben die Web-Seiten von Gary King.

Open Data aus Sicht von APA, DGPs und BDP

Die [Ethical Principles of Psychologists and Code Of Conduct](#) der APA (2002) fordern im Absatz 8.14 Sharing Research Data for Verification:

- (a) After research results are published, psychologists do not withhold the data on which their conclusions are based from other competent professionals who seek to verify the substantive claims through reanalysis and who intend to use such data only for that purpose, provided that the confidentiality of the participants can be protected and unless legal rights concerning proprietary data preclude their release. This does not preclude psychologists from requiring that such individuals or groups be responsible for costs associated with the provision of such information.
- (b) Psychologists who request data from other psychologists to verify the substantive claims through reanalysis may use shared data only for the declared purpose. Requesting psychologists obtain prior written agreement for all other uses of the data.

Die von der APA formulierten Prinzipien für [Data Acquisition, Management, Sharing and Ownership](#) lehnen sich an die Prinzipien der *National Science Foundation* an. Diese Internet-Seite der [APA](#) gibt außerdem Links zu “US Federal policies and non-governmental organizations involved in data sharing and archiving”. Die Seiten [Access to and Retention of Research Data: Rights and Responsibilities](#) des Council on Governmental Relations COGR, einer Vereinigung von 150 Universitäten der USA, enthalten zahlreiche Links zu Informationsquellen, z.B. zur [NIH Data Sharing Policy](#) von 2003, die für alle vom National Health geförderten Forschungsarbeiten (mit einer Bewilligungssumme > \$500.000) verlangt, dass diese „must have a plan to address the sharing and archiving of data“. Weitere Links führen zu typischen data repositories and portal sites, u.a. zu Gesundheitsforschung, zu bekannten epidemiologischen Studien u.a. Klinischen Studien, zu Datensätzen berühmter Längsschnittstudien aus der Entwicklungspsychologie, Schul- und Bildungsstatistiken, z.T. mit internationaler Datenbasis, zu einem Brain Imaging (fMRI)-Archiv usw.

Die [Deutsche Gesellschaft für Psychologie e.V.](#) und der [Berufsverband Deutscher Psychologinnen und Psychologen e.V.](#) haben *Ethische Richtlinien* publiziert, die auch Hinweise auf den Umgang mit Primärdaten wissenschaftlicher Arbeit geben. Die Ethischen Richtlinien besagen – unter Bezug auf die Richtlinien der APA – im Abschnitt

C.III. Grundsätze für Forschung und Publikation

14. Weitergabe von Forschungsdaten zum Zweck der Überprüfung

- (a) Nach der Publikation von Daten halten Psychologen die Originaldaten nicht zurück, wenn andere Wissenschaftler und Wissenschaftlerinnen ihre Schlussfolgerungen durch Re-Analyse überprüfen wollen und die Daten nur zu diesem Zweck nutzen wollen. Dies gilt allerdings nur in solchen Fällen, in denen die Weitergabe der Daten erfolgen kann, ohne dass die Vertraulichkeit personenbezogener Informationen gefährdet ist, und falls keine gesetzlich geregelten Eigentumsrechte der Herausgabe entgegen stehen. Dies impliziert keine kostenlose Weitergabe von Daten; für die Weitergabe der dokumentierten Daten kann der entsprechende Aufwand in Rechnung gestellt werden.

(b) Psychologen, die von Kollegen und Kolleginnen Daten zur Re-Analyse mit dem Ziel der Überprüfung der Schlussfolgerungen anfordern, dürfen diese nur zu dem in der Anfrage angegebenen Zweck verwenden. Für jeden anderweitigen Gebrauch der Daten ist eine vorherige schriftliche Genehmigung einzuholen.“

(Ethische Richtlinien der Deutschen Gesellschaft für Psychologie e.V., revidierte Fassung vom 28.9.2004, und des Berufsverbands Deutscher Psychologinnen und Psychologen e.V., Fassung vom 16.4.2005).

Einige allgemeine Informationen gibt die von mehreren deutschen Universitätsbibliotheken geschaffene [Open Access Informationsplattform](#) auf den Seiten der Psychologie.

Motive der Open Access Bewegung

Von den Befürwortern eines geregelten Zugangs zu Forschungsdaten zwecks Vergleich, Qualitätssicherung, Reanalyse und Metaanalyse werden wissenschaftsinterne und auch gesellschaftliche bzw. ökonomische Gründe genannt. So zählten Wicherts, Borsboom, Kats und Molenaar (2006) sechs Aspekte des wissenschaftlichen Nutzens auf:

“First, it would become much easier to reanalyze data; one could, for instance, quickly and effectively confirm the robustness of results with respect to various statistical analyses (e.g., parametric or nonparametric techniques). Second, the feasibility of doing meta-analyses would be greatly enhanced, which would facilitate assessments of the magnitude of effects and the stability of research findings. Third, it would be easy to address any question one might entertain with respect to data massaging or overly favorable descriptions of results; in addition, such instances of dubious science would, in all likelihood, become highly infrequent as soon as authors knew that other researchers would be looking over their shoulders. Fourth, the procedure would open up a new critical dimension in science because it would be not just the researcher’s interpretation of the data that entered the public discourse of science, but the data themselves. Fifth, once data were available on the Web, they would remain available (to all, including the author himself or herself), so that the scientific process would no longer be hampered by the fact that, after a few years have passed, researchers are often unable to locate their data, may have moved to a different university, or may have discontinued scientific work altogether. Sixth, and perhaps most important, scientific evidence should be publicly accessible as a matter of principle; anybody who wants to play in the scientific arena will have to come in with open sight. A procedure like the one proposed would thus increase the openness of scientific research. It seems to us that, considering the ratio of the benefits achieved in this manner to the costs involved in terms of extra work, this is a bargain. We therefore suggest that the APA journals incorporate the proposed procedure in the publication process. (p. 727-728).

Außer diesen Perspektiven der Reanalyse und methodischen Qualitätskontrolle werden zunehmend die gesellschaftlichen und ökonomischen Motive für Open Access Forschungsdaten genannt. Insbesondere wenn die Forschungsarbeit aus öffentlichen Haushalten, d.h. durch Universitäten, die Deutsche Forschungsgemeinschaft, Max-Planck-Gesellschaft, Leibniz-Gesellschaft u.a., finanziert wird, lassen sich Forderungen nach ökonomischer Verwendung der Steuergelder sowie nach Zugänglichkeit für andere Wissenschaftler zwecks Partizipation und kritischer Reanalyse ableiten. Ein effektives Data Sharing könnte beitragen, die Kosten zu reduzieren bzw. die Mittel ökonomischer einzusetzen. Zum Thema „Wissenschaft als öffentliches Gut“ und „Information als gesellschaftliche Ressource“ bzw. über die Ungleichverteilung des Wissens und des Informationszugangs finden sich Argumente u.a. bei Herb (2006) und Mruck, Gradmann und Mey (2004). Fröhlich (2009) wies darauf hin, dass Open Access ganz im Sinne der wissenschaftstheoretischen Auffassungen von Karl Popper und Robert Merton liegen würde.

Andererseits haben Wissenschaftler, die ihren Forschungsvorsprung sehen oder an Patente und Lizenzen denken, berechnete Interessen, Forschungsdaten – zunächst – für sich zu behalten. Auch Klinische Einrichtungen, Personalassessment-Center, Kranken- und Rentenversicherungen können motiviert sein, ihre für die Forschung hochinteressanten Primärdaten *nicht* zugänglich zu machen. – Die zu schützenden Bereiche und Datenarten sollten jedoch möglichst genau abgegrenzt werden. Die 2003 publizierte Forderung des National Institutes of Health (NIH) bei allen Projekten (mit direkten Kosten höher als \$500,000 jährlich) bereits im Antrag, die geplante Daten-Archivierung zu spezifizieren, verlangt frühzeitige Entscheidungen, wie Zustimmung und Vertraulichkeit gewahrt, das Data Sharing ermöglicht bzw. Abgrenzungen begründet werden können.

Weitere Bedenken gegen die sekundäre Nutzung von Daten betreffen u.a. Sorgen über den möglichen Missbrauch oder die mögliche Verletzung der Anonymität (siehe u.a. Johnson & Sabourin, 2001; Law, 2005). Die Anonymisierung „sensibler“ Daten ist selbstverständlich. Darüber hinaus sind die sonstigen Missbrauchsmöglichkeiten meist ziemlich vage formuliert. Genannt werden u.a. die unberechtigte Weitergabe der Daten oder die unberechtigte Nutzung (eventuell auch Veränderung/Verfälschung). Solche Vorgänge werden jedoch auch in den Forschungsdatenbanken unter das Urheberrecht fallen, so dass es sich eher um die Frage der Durchsetzung (Forderung nach Unterlassung, Schadenersatz

usw.) handeln dürfte. Die Verträge mit den Datennehmern halten diese grundsätzlichen Vereinbarungen fest (zur juristischen Seite in Deutschland, siehe Spindler, 2006).

Schließlich wird auf die Schwierigkeiten der praktischen Nutzung, auf die wünschenswerte Harmonisierung von Systemarchitektur und Protokollen, auf den organisatorischen Aufwand und die Kosten hingewiesen. Auf der anderen Seite existieren bereits wichtige Forschungsdatenbanken, die unentgeltlich genutzt werden und unentbehrlich geworden sind.

Zwischen den drei typischen Positionen „Halte Deine Daten geheim“, „Keep your data and share them if requested“ und „Daten sind ein unverzichtbarer Teil der wissenschaftlichen Publikation“ gibt es offensichtlich ein breites Spektrum der Auffassungen.

Reanalysen und Metaanalysen

„Data Sharing“ als Teilhabe an Daten zur Nachnutzung kann als ein konstitutives Prinzip wissenschaftlicher Arbeit gelten, denn – auf lange Sicht und im Prinzip auf Gegenseitigkeit – ist Forschung auf die Kumulation und Aggregation der Daten zur überzeugenden Hypothesenprüfung angewiesen. Aber auch unmittelbar können Reanalysen im Interesse der Autoren liegen, wenn z.B. umfangreiche Datensätze nicht vollständig ausgewertet sind. So haben große sozialwissenschaftliche Umfragen, aufwendige psychologische Längsschnittstudien und große Therapiestudien häufig einen kaum auszuschöpfenden Informationsgehalt. Auch in der differentiellen und klinischen Psychologie existieren, z.B. in vielen Projekten mit *Ambulantem Assessment*, mehrdimensionale (multi-level) Datensätze (Personen, Situationen, Zeitpunkte, primäre und aggregierte Variablen). Solche Datensätze haben wegen der verschiedenen Datenebenen, Perspektiven und Aggregationsmöglichkeiten außer der primären Fragestellung noch vielfältige Auswertungschancen.

Die von vielen Reviewern zu zentralen Fragestellungen in verschiedensten Bereichen der Psychologie berichtete und beklagte Inkonsistenz der Untersuchungsbefunde hat früher zur Forderung nach identischer Replikation innerhalb und zwischen den Instituten geführt (Eid & Diener, 2006; Schweizer, 1989). Gegen dieses in den Naturwissenschaften gängige *Kontrollprinzip* scheint es in der psychologischen Forschung und bei vielen Herausgebern deutliche Reserven zu geben. – Eventuell könnten die auf der *Ebene von Primärdaten* durchgeführten Reanalysen und Metaanalysen einen teilweisen Ersatz schaffen. Oder würden sie eventuell zunächst noch tiefer in die Methodenprobleme führen: u.a. unzureichende Parameterdefinitionen, mangelnde Explikation der experimentellen Rahmenbedingungen und anderer Kontexte, Fehlen von Konventionen über Methodenstandards?

Data Sharing, Data Recycling, Data Mining – die Begriffe existieren, die reale Ausführung, zumindest in der Psychologie, dürfte vorerst schwierig sein, weil der Zugang zur Datenbasis noch weitgehend fehlt.

Forschungsdatenbanken

Eine ausführliche Übersicht über Forschungsdatenbanken (nicht Dokumentenserver für Texte) der Sozialwissenschaften und der Psychologie, etwa für den deutschsprachigen Bereich, ist hier nicht beabsichtigt (siehe die zitierten Links). Herausragende *Beispiele* sind:

- das ALLBUS-Surveyprogramm, in dem seit 1980 regelmäßig im Abstand von zwei Jahren repräsentative Querschnitt-Stichproben der Bevölkerung befragt wurden: zur sozialen Lage und zum sozialen Wandel in Deutschland, zu Einstellungen, Werten und Verhaltensweisen, z.T. mit Fragen zu sozialen und weltanschaulichen Einstellungen, Wertvorstellungen, Verhaltensweisen und Persönlichkeitsmerkmalen, mit dem Träger [GESIS - Leibniz-Institut für Sozialwissenschaften](#);
- Internationale Erhebungen, die eine Vielzahl von Ländern und relativ lange Zeiträume abdecken, u.a. Eurobarometer ([Eurobarometer bei GESIS](#)), [European Values Study – EVS \(EVS bei GESIS\)](#), [International Social Survey Programme – ISSP \(ISSP bei GESIS\)](#);
- das [Sozioökonomische Panel SOEP](#) des Deutschen Institut für Wirtschaftsforschung, Berlin.

Im Bereich der „qualitativen“ Methoden sind erwähnenswert:

- [QUALIDATA](#), das *Qualitative Data Archival Resource Centre* des englischen *Economic and Social Science Research Council* (ESRC) (siehe Corti (2000));
- [ALLF Archiv für Lebenslaufforschung](#).

Dagegen sind in dem innovativen und umfangreichen [Forum Qualitative Sozialforschung FQS](#) nur Ergebnisberichte und andere sekundäre Texte und bisher keine Materialien oder numerischen Datensätze vorgesehen (vgl. jedoch entsprechende Ansätze in der aktuellen Diskussionen auf den Open-Access-Tagungen des FQS und assoziierter Institutionen). Der Begriff *Daten* meint hier wesentlich mehr als – im engeren Sinn – nur numerische Daten. Die empirische

Basis psychologischer Untersuchungen umfasst u.a. primäre Textquellen, z.B. Protokolle von Interviews oder Beobachtungen, Tagebücher, Biografien, narratives Material, Alltagsgeschehen, Audio- und Video-Aufzeichnungen verschiedenster Art.

In Deutschland gibt es seit 2002 das Vorhaben *DPI - Digitale Psychologie Information: Ein Gemeinschaftsprojekt der Saarländischen Universitäts- und Landesbibliothek (SULB) und des Zentrums für Psychologische Information und Dokumentation (ZPID)*. Zu nennen sind hier die [Virtuelle Fachbibliothek Psychologie](#) und PsyDok, der Psychologie Volltextserver, auf dem Internetpublikationen dauerhaft archiviert sind.

Was die Archivierung von Primärdaten und das entsprechende Data Sharing zur Nachnutzung in der Psychologie angeht, ergibt sich der Eindruck, dass zwar die Prinzipien bejaht werden, sich jedoch im Vergleich zu einigen der Naturwissenschaften noch relativ wenig geändert hat. – Andererseits besteht in Deutschland bereits seit 2002 im Zentrum für Psychologische Information und Dokumentation ZPID die zentrale Einrichtung PsychData.

ZPID PsychData

Die Ziele des aus Mitteln der DFG gegründeten psychologischen Datenarchivs [PsychData](#) des [ZPID](#) werden auf der Homepage beschrieben: „die umfassende Dokumentation und Langzeitarchivierung von Primärdaten mit Hilfe – eigens dafür zusammengestellter – Metadaten aus allen Bereichen der psychologischen Forschung und sonstiger Humanwissenschaften sowie die Bereitstellung dieser Datensätze für wissenschaftliche Nutzungszwecke wie Sekundär- und Reanalysen.“

„Die Regeln der Deutschen Forschungsgemeinschaft (DFG) zur Sicherung guter wissenschaftlicher Praxis (19981) werden als allgemein akzeptierte Grundhaltung von allen Beteiligten anerkannt. „Als Archiv gewährleisten wir die plattformunabhängige, zerfallsfreie und langfristige Aufbewahrung und Bereitstellung der uns anvertrauten Dokumente. Ressourcen, die durch PsychData zur Verfügung gestellt werden, sind von der kommerziellen Nachnutzung ausgeschlossen.“

„Ziel des Psychologischen Datenarchivs PsychData ist die Bereitstellung der dokumentierten Datensätze für Sekundäranalysen verschiedener Art:

- die Analyse der Datensätze unter neuen Fragestellungen
- die Reanalyse der Datensätze mit Hilfe alternativer methodischer Zugänge
- die vergleichende Analyse verschiedener Datensätze
- die Aggregation von Datensätzen zur Optimierung von Parameterschätzungen
- die Ermittlung historischer Veränderungen

Das ZPID möchte Ihnen als empirisch Forschenden mit dem Psychologischen Datenarchiv PsychData ein Serviceangebot unterbreiten das folgende Richtlinien hat:

- Schutz Ihrer Datensätze vor Datenverlust, -alterung und -manipulation
- Nutzung Ihrer Datensätze gemäß vertraglicher Vereinbarungen
- Erfüllung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis
- Unterstützung der Rezeption Ihrer Forschungen durch die Fachöffentlichkeit
- Ausweis Ihrer Datensätze als elektronische Publikationen
- Vermeidung redundanter Datenerhebungen und
- Verstärkung des wissenschaftlichen Austauschs

Das Gelingen des Vorhabens hängt wesentlich von Ihrer Bereitschaft ab, durch die Freigabe eigener Datensätze an dem Aufbau des Archivs aktiv mitzuwirken. Dabei entstehende Unkosten können durch die DFG erstattet werden.“

Bei PsychData ist an die gesamte Psychologie gedacht: „Dabei ist die Aufnahme von Datensätzen aus der psychologischen Umfragemforschung und experimentellen Psychologie ebenso möglich wie die Archivierung von Datensätzen, die im Zuge von Testentwicklungen erhoben wurden.

Ziel des Psychologischen Datenarchivs PsychData ist die Bereitstellung der dokumentierten Datensätze für Sekundäranalysen verschiedener Art.“

Aktuell sind 33 Studien archiviert, davon 7 experimentelle Studien, vorwiegend aus der Allgemeinen Psychologie, 23 hauptsächlich auf Fragebogen gestützte Untersuchungen und Längsschnittstudien, vorwiegend sozial-, gesundheits- und entwicklungspsychologischer Art, sowie 3 Datensätze psychologischer Testkonstruktionen. Dazu gehört die Normierungsstichprobe des von Becker konstruierten Trierer Persönlichkeitsfragebogen TPF (nicht zu verwechseln mit dem

neueren TPI), der Mehrdimensionale Befindlichkeitsfragebogens MDBF von Steyer, Schwenkmezger, Notz und Eid sowie die von Krampen konstruierten Skalen zur Erfassung von Hoffnungslosigkeit.

Zwischen dem ZPID und dem *Datengeber* wird ein Nutzungsüberlassungsvertrag und zwischen ZPID und *Datennnehmer* ein Nutzungsvertrag geschlossen, um die wechselseitigen Verpflichtungen zu regeln. – Demgegenüber wird z.B. von GESIS – mit Hinweisen auf die allgemeinen Urheberrechte – nur eine kurze Registrierung der Adresse, der Institution und der Nutzungsabsichten verlangt.

Der aktuellen Übersicht von Ruusalepp (2008) zufolge kann die ZPID PsychData-Einrichtung international als der einzige Fall eines digitalen Psychologischen Datenarchivs gelten. Es ist ein hervorragendes *disziplinäres* Repositorium für Primärdaten, wird aber offensichtlich noch viel zu wenig genutzt.

Zeitschriften-Publikationen – nur noch mit den Daten?

Die ethischen Richtlinien der APA über wissenschaftliche Publikationen lassen bisher offen, wo und wie die primären Datensätze dokumentiert und archiviert werden sollen. Die APA-Journals haben dieses Prinzip bisher nicht umgesetzt. Über die Gründe lassen sich nur Vermutungen anstellen. Es scheint auch keine zeitliche Perspektive für die Implementierung einer Data Archiving Policy zu geben.¹

Auf den zitierten Abschnitt der Ethical Standards (bzw. des Publication Manual) der APA bezogen sich Wicherts et al. (2006) als sie die Autoren von Artikeln in vier verschiedenen APA-Zeitschriften um die Überlassung der Datensätze zum Zwecke einer Reanalyse über die Bedeutung von Ausreißer-Werten baten. Zu den 141 ausgewählten Artikeln erhielten sie nur in 27 % der Fälle tatsächlich die Daten; sie mussten daraufhin ihr Vorhaben abbrechen. Die Sachlage scheint sich in den vier Jahrzehnten seit Wolins (1962) frustrierenden Erfahrungen nicht wesentlich geändert zu haben. Deswegen schlussfolgerten Wicherts et al. (2006):

“Only two modifications are necessary, namely: (a) that upon the acceptance of a manuscript for publication, authors be required to submit an ASCII file with anonymized data, as well as a codebook in standardized format and (b) that the journal publish the data and codebook on the World Wide Web as an electronic appendix to the article.” (Wicherts et al., 2006, p. 727). Gemeint sind hier nicht die kompletten Forschungsdaten, sondern nur die in der Publikation verwendeten Daten.

Viele Wissenschaftler würden wahrscheinlich die vorgeschlagene Regelung, vielleicht zunächst auf freiwilliger Basis, später obligatorisch, sehr begrüßen, weil dies einen großen Fortschritt für Reanalysen und Qualitätssicherung bedeutete.

Ungeklärte Rechtslage

Die gegenwärtige Diskussion über Open Access bei wissenschaftlichen Text-Publikationen ist überlagert von der Auseinandersetzung über verwandte Themen: die sog. Retrodigitalisierung älterer, nicht mehr nachgedruckter Bücher und die Wahrung von Urheberrechten. In der von zahlreichen Verlegern und Autoren unterzeichneten Erklärung [Für Publikationsfreiheit und die Wahrung der Urheberrechte](#) wird behauptet, dass durch die beabsichtigten Open Access Regelungen Grundrechte verletzt werden.

Dieser Vorwurf führte zu einer [Gemeinsamen Erklärung der Wissenschaftsorganisationen. Open Access und Urheberrecht: Kein Eingriff in die Publikationsfreiheit](#) (25. März 2009) mit dem Hinweis, dass in dieser Diskussion zu wenig zwischen den verschiedenen Sachverhalten und Interessen unterschieden werde. Zu dieser Allianz gehören u.a. Deutsche Forschungsgemeinschaft, Fraunhofer Gesellschaft, Helmholtz-Gemeinschaft Deutscher Forschungszentren, Hochschulrektorenkonferenz, Leibniz-Gemeinschaft, Max-Planck-Gesellschaft, Wissenschaftsrat.

„In der Erklärung „Für Publikationsfreiheit und die Wahrung der Urheberrechte“, die von zahlreichen Verlegern und Autoren unterzeichnet wurde, sieht sich die Allianz der deutschen Wissenschaftsorganisationen mit der inakzeptablen Unterstellung konfrontiert, die Freiheit zur Veröffentlichung in grundgesetzwidriger Weise beschneiden zu wollen.

1. Die Allianz der Wissenschaftsorganisationen fordert eine für den Leser entgeltfreie Publikation (Open Access) ausschließlich von Forschungsergebnissen, die durch den Einsatz öffentlicher Mittel und damit zum Nutzen der Forschung und Gesellschaft insgesamt erarbeitet wurden. Keinesfalls fordert die Allianz eine Open-Access-Publikation

¹ Anfragen beim Senior Director APA Journals und bei einzelnen Herausgebern führten noch zu keiner konkreten Information über die Absichten bzw. den Stand der Implementierung hinsichtlich data-enriched publications, archiving and sharing of primary research data.

belletristischer Schriften, aus deren Verwertung Autoren ihren Lebensunterhalt beziehen. Dies zu suggerieren, ist irreführend.

2. Die Allianz der Wissenschaftsorganisationen fordert nicht zum Rechtsbruch auf. Die Open-Access-Politik der Allianz greift weder in die Urheberrechte der Autoren ein, noch lässt sie die Verlagsinteressen unberücksichtigt. Im Gegenteil: Die Förderung der Umstellung des wissenschaftlichen Publikationswesens geht mit der Umschichtung von Mitteln einher, die neue, Open-Access-kompatible Geschäftsmodelle ermöglichen. Diese Angebote werden von Verlagen aufgegriffen. Open Access ist für viele ein funktionierendes Geschäftsmodell. Alle Richtlinien, mit welchen Wissenschaftlerinnen und Wissenschaftler in Deutschland oder in anderen Staaten zu Open-Access-kompatiblen Publizieren aufgefordert werden, verknüpfen das Open-Access-Publizieren mit dem geltenden Urheberrecht. Der Vorwurf einer „Enteignung der Urheber“ entbehrt jeder Grundlage, denn Wissenschaftlerinnen und Wissenschaftler bleiben nach wie vor alleinige Urheber ihrer Werke.

3. Die Freiheit der Wissenschaft ist ein hohes und im Grundgesetz verbrieftes Gut, das konstitutives Merkmal der Allianzorganisationen ist. Die bei uns beschäftigten oder von uns geförderten Wissenschaftlerinnen und Wissenschaftler sind frei in der Wahl ihrer primären Publikationsformen. Wir erwarten jedoch, dass die Autoren der Gesellschaft, die ihre Forschung durch Steuermittel möglich macht, einen einfachen Zugang zu ihren Publikationen eröffnen, der zudem die öffentliche Hand möglichst wenig finanziell belastet. Wo dies primär nicht durch die Veröffentlichung in einem Open-Access-Medium geschehen kann, sollen die Autoren ihre Publikationen über Open-Access-Repositorien verfügbar machen, wo immer dies rechtlich möglich ist. Der Großteil gerade der international agierenden Verlage gestattet dies schon heute.

4. Soweit öffentlich geförderte Forschung betroffen ist, bestreiten die Partner der Allianz mit ihrem finanziellen Engagement die gesamte Kette der Wissensverwertung von der Informationsbeschaffung über die erste Lektüre und die ersten Laborversuche bis hin zum Kauf der mit öffentlichen Mitteln erstellten Publikationen. Sofern Verlagsdienstleistungen in Anspruch genommen werden, um primär im Open Access zu publizieren, sind diese selbstverständlich nach wie vor zu vergüten. Genau hierfür stellen die Allianz-Organisationen einen Teil ihres Budgets bereit.

Die zitierten Abgrenzungen und die Hervorhebung öffentlich geförderter Forschungsvorhaben sind wichtig (siehe auch Kleiner, 2009), doch gibt es schwierige Grenzbereiche und insgesamt ein Spannungsfeld unterschiedlicher rechtlicher Aspekte.

Wie unzureichend die rechtlichen Grundlagen geklärt sind, lassen die Vortragsankündigungen für die im November 2009 stattfindende Urheberrechtstagung <http://www.urheberrechtstagung.de/> erkennen. Mit der Entwicklung zu einer web-basierten Informationsgesellschaft reichen die Begriffe des bisherigen Urheberrechts nicht mehr aus, so dass tiefreichende Modernisierungen unumgänglich sein werden. So können Eigentumsrechte, Urheberrechte, Nutzerrechte, verlegerisches Copyright, Zugangsrechte unterschieden und abgewogen werden. Hierbei scheint unklar zu sein, wie Eigentumsrechte (an sachlichem bzw. geistigem Eigentum) an Primärdaten der Forschung geltend gemacht werden können. Für welche Kategorien von Primärdaten sind *besondere Regelungen* (siehe oben) zu treffen? Eine zentrale Frage ist: Rechtfertigt die Forschungsförderung aus öffentlichen Mitteln eine Steuerung des Publikationsverhaltens?

Wenn bereits die Rechtslage in Deutschland unklar ist, so wird sie international durch die großen Unterschiede der Rechtsvorschriften anderer Länder weiter kompliziert. Es heißt, dass die DFG und die Europäische Kommission durch Gutachten die juristischen Rahmenbedingungen klären lassen wollen, um dann vielleicht eine Vereinheitlichung zu erreichen.

Psychologische Tests

Die Datenbasis psychologischer Tests zugänglich zu machen, sollte ein Kernthema der Qualitätssicherung sein. Hier muss jedoch unterschieden werden: Die *Teststatistiken* und die *Normen* (auf der Ebene der Testwerte) sind ohnehin im Test-Manual veröffentlicht. Demgegenüber sind die *Normierungsdaten auf der Ebene der Items*, wie sie für die Testkonstruktion verwendet wurden, in der Regel unzugänglich. Eine Weitergabe an Kollegen zwecks Reanalyse oder Vergleich ist in Einzelfällen nicht ganz unüblich, unterscheidet sich jedoch grundsätzlich von einer breiten Open Access Regelung.

Die Fachverlage sind Wirtschaftsunternehmen und könnten bei open-access-Regelungen einen Missbrauch befürchten; auch die Testautoren könnten sich entsprechende Sorgen machen. Doch wie könnte überhaupt ein Missbrauch der Daten konkret geschehen? Selbst wenn es aufgrund des Open Access gelingen sollte, aus den primären Item-Verteilungen einen sehr ähnlichen oder einen testmethodisch verbesserten Test abzuleiten, würden ja der Name des Tests und das Copyright des Originals einschließlich des Testmaterials weiterhin unter urheberrechtlichem Schutz stehen.

Die Testautoren haben zwar in der Regel mit der Publikation des Tests das Copyright für ihren Test (und für etwaige Lizenzen) an den Verlag übertragen, aber den Testautoren bleiben – neben der ausschließlichen wissenschaftlichen

Verantwortung – die Rechte an ihren primären Daten, die eigens für die Testkonstruktion erhoben wurden. Die bisherigen Verlagsverträge haben diesen Sachverhalt nicht ausreichend präzisiert. – Auch bei der Publikation eines wissenschaftlichen Artikels gehen die zugrundeliegenden *Primärdaten* keinesfalls in das Eigentum und Copyright des Zeitschriften-Verlags über.

Open Access zu den Testentwicklungsdaten spielt im gegenwärtigen *Testbeurteilungssystem des Testkuratoriums TBS-TK* (2007) der Föderation Deutscher Psychologinnenvereinigungen bzw. in der DIN 33430 noch keine Rolle. Eine gründliche Qualitätskontrolle ist jedoch nur anhand der primären Daten möglich. Das TBS-TK kann zwar nicht die Dokumentation der Primärdaten regeln, jedoch eventuell Empfehlungen geben, wie diese Aufgaben als Vorbedingung vertiefter Qualitätskontrolle zu lösen wären. Wo und wie sind die primären Daten dokumentiert und gesichert? Ist ein Zugang zum Zweck von Reanalysen möglich? Wenn die primären Daten gewöhnlich nur bei den Testautoren liegen, stellt sich – zumindest bei den häufig und längerfristig verwendeten Tests – die Frage, wie solche Basisdaten über die Zeit des aktiven Engagements der Testautoren hinaus gesichert sind.

Auch die Testautoren könnten durchaus an Reanalysen interessiert sein, z.B. wenn nach klassischer Testtheorie konstruierte Intelligenz- und Leistungs-Tests, eventuell auch Persönlichkeitsskalen, nach probabilistischen Konzepten rekonstruiert werden, um Übereinstimmungen und Widersprüche zu erkennen.

Recherchen im Internet führen gelegentlich zu Seiten mit „normative data“ psychologischer Tests, doch handelt es sich in den allermeisten Fällen um Vergleichsdaten, d.h. *Testwerte* bestimmter Untersuchungsgruppen und nicht um die zur *Testkonstruktion* verwendete *Datenbasis*.

Ein seltener erwähnter Vorzug des open-access ist, dass geeignete Datensätze für Ausbildungszwecke zu erhalten wären. So mangelt es z.B. den meisten Lehrbücher der Testmethodik an Daten bekannter und häufig verwendeter Tests, die – ausgewählt und kommentiert – viel interessanter wären als die üblichen, mageren oder artifiziellen Datensätze. In der wissenschaftlichen Ausbildung, in den Kursen zur Methodenlehre und anderen Lehrveranstaltungen könnten in vielen Fällen didaktisch nutzbringend wirkliche Datensätze für Übungszwecke herangezogen werden, allerdings stellen sich bei dieser Weitergabe an mehrere Personen zusätzliche rechtliche Fragen, welche die spezielle Zustimmung des Autors (und ggf. des Verlags) erfordern.

Ausblick: Open Data in der Psychologie

Innerhalb der Psychologie geht es gegenwärtig fast nur um Open Access Regelungen für *Texte* verschiedener Art. Auch die *Primärdaten* wichtiger psychologischer Untersuchungen, herausragender Experimente, Testkonstruktionen, Umfragen usw. einheitlich zu dokumentieren, sicher zu archivieren und zugänglich zu machen, bedeutete zweifellos einen Fortschritt der wissenschaftlichen Arbeit in vielerlei Hinsicht. Dazu gehören vergleichende Beurteilungen, eine Vertiefung der kritischen Reanalysen, gründlichere Metaanalysen, Qualitätssicherung und ggf. neue didaktische Möglichkeiten.

Die hier skizzierten Recherchen ergaben ein frühes Interesse am öffentlichen Zugang zu Primärdaten der Psychologie, dann einen relativen Stillstand und mit der Entwicklung der Internet-Kommunikation neue Initiativen der Datenarchivierung – in der Praxis oft noch zögerlich und im Bereich der psychologischen Tests unterentwickelt.

Die erst in Umrissen zu erkennenden vielfältigen Bedingungen und Probleme müssten von den Experten kommentiert werden. Diese Experten-Diskussion ist schon seit längerem im Gange und wird einmal zu weitreichenden Konsequenzen führen. Hier werden nur vier allgemeinere Aspekte skizziert.

Bewusstseinsänderung

PsychData ist die 2002 für diese Archivierung geschaffene Einrichtung des ZPID. Offensichtlich wird sie erst vereinzelt genutzt. Andererseits würden verstärkte Appelle eine häufigere Inanspruchnahme zu Folge haben, was angesichts der notwendigen und anerkanntenswert sorgfältigen Hilfestellung einen vermehrten Arbeitsaufwand für das ZPID bedeutet. Damit werden Finanzierungskonzepte und zusätzliche Projektmitarbeiter notwendig. Die ausdrückliche Öffnung von PsychData für alle Teilbereiche der Psychologie ist wünschenswert.

Offene Daten in der Wissenschaft sind gegenwärtig für viele noch eine befremdliche Vorstellung. Die Diskussion über die Motive, Vorzüge und Grenzen werden, wie gegenwärtig bei den e-Journals und der Open Access Archivierung von Texten auf Dokumentenservern, zunehmen. Dass sich die Einstellungen ändern werden, ist angesichts der sachlichen Gründe und nicht zuletzt wegen der attraktiven Internet-Möglichkeiten sehr wahrscheinlich. Vermutlich hat PsychData noch keinen der Bedeutung angemessenen Bekanntheitsgrad in der Fachwelt. Vielleicht wird es einmal eine spezielle

Umfrage geben, analog jener der DFG zu Open Access Publikationen. Mehr Aufmerksamkeit für diese wohl unumgängliche Entwicklung könnte einerseits durch verstärkte Hinweise auf PsychData in der Fachöffentlichkeit und andererseits durch Hervorhebung beispielhafter Datensätze geschehen, wie es vom ZPID vorgesehen ist. Weiterhin könnte die Archivierung bestimmter Studien aktiv angeregt werden: als hervorragend beurteilte, innovative, von Experten

nominierte, mit Wissenschaftspreisen ausgezeichnete Studien, wobei auch quantitative Merkmale wie Impactfaktoren und die Kostenschätzungen für etwaige Replikationen einbezogen werden könnten. Diese Präzedenzfälle könnten die Einstellungsänderung und die künftige, systematische Archivierung fördern. – Für wie wichtig oder unwichtig wird die Dokumentation und Nachnutzung von Primärdaten im Selbstverständnis des Faches Psychologie künftig gehalten werden?

Antizipation des Aufwands

Sind die von NSF und NIH vertretenen Prinzipien und die von Wicherts et al. (2006) für die Fachzeitschriften erweiterte Konzeption realistisch, die relevanten Primärdaten systematisch (1) zu allen mit öffentlichen Mitteln geförderten Forschungsarbeiten (oberhalb einer festzulegenden Größenordnung?) und (2) zu jeder qualifizierten Publikation zu archivieren und zugänglich zu machen? Diese Vorstellungen wären, trotz der zu antizipierenden Größenordnung, auch im deutschsprachigen Raum technisch umsetzbar. Nicht die Datenserver, sondern der Arbeitsaufwand und die Tendenz zur Geheimhaltung setzen die Grenzen.

Was ist für die deutschen Verhältnisse zu antizipieren? Zwar bezieht sich in der Psychologie nur ein Teil der Zeitschriften-Publikationen auf einen eigenen Datensatz, doch wäre zusammen mit den öffentlich geförderten Projekten (mit Beträgen > 300.00 Euro?) jährlich eine dreistellige Zahl von Datensätzen zentral oder dezentral zu bewältigen. – Je nach Einstellung und eigener Erfahrung: ein überbordender Datenfriedhof oder eine neue und wissenschaftlich produktive Ära des Data Sharing, Data Recycling, Data Mining.

Standardisierungen und Formatierungen

Aus Sicht der Forschungsdatenbanken ist die hochgradige Standardisierung der Dokumentation erstrebenswert (vgl. [Data Documentation Initiative](#)). Ein praktisch wichtiges und problematisches Detail ist die Vereinheitlichung des Datenformats. Im Interesse der dauerhaften Archivierung scheint der einfache *American Standard Code for Information Interchange* ASCII für Textdateien (.txt) zweckmäßig zu sein, um die „Migration“ in künftige Systeme zu erleichtern. Die Datengeber liefern die Primärdaten in Form einer tabulatorgetrennten Textdatei oder eines gleichwertigen Formats. Dass ZPID betont, dass die PsychData-Standards nicht identisch mit denen anderer Archive sind, der verwendete Metadatensatz sei jedoch auf die meisten anderen Standards abzubilden. (Allerdings ist abzusehen, dass die Methoden der Textanalyse, die Videoanalyse, die Dokumentation sonstiger Materialien ohnehin spezielle Formate und Dokumentationsweisen erfordern.) Demgegenüber ermöglicht die GESIS-Forschungsdatenbank ein Download der ALLBUS-Dateien als SPSS file (.sav Format) und der Datendokumentation einschließlich des Methodenberichts im .pdf Format.²

Die meisten Autoren im Fach Psychologie werden ihre Datensätze in Statistiksystemen, zumeist wohl im SPSS (sowie SAS oder anderen) angelegt haben. Nun müssen für .txt-Archivierung die *Daten* von den *Variableninformationen* (Name, Kurzbeschreibung mit Wertedefinition) getrennt und letztere in einem Kodebuch dokumentiert werden. Bei größeren oder bei multi-level Datensätzen kann das einigermaßen umständlich sein. Andererseits wird ein solches Kodebuch zur erschöpfenden Beschreibung der Methodik selten ausreichen. Von Fragebogendaten abgesehen sind viele oder die meisten der heutigen Untersuchungen so beschaffen, dass zu dem Datensatz unbedingt die Publikationen, die genaueren Methoden-Beschreibungen sowie weitere Quellen verwendet werden müssen. Beispiele sind die psychologische Struktur einer Beobachtungssituation, die biologisch-physiologischen Messtechniken (Biosignale, Hormon-Assays), spezielle Rahmenbedingungen eines Experiments, Kontexte und Ablauf eines Interviews, spezielle Software für Datenaufnahme und Auswertungstechniken. Die Autoren sollten – eher als in die Formatierungen – mehr Zeit in die möglichst ausführliche Studienbeschreibung investieren, vor allem für die in den Zeitschriftenartikeln meist viel zu kurz kommenden methodischen Standards, Konventionen und ggf. Besonderheiten – auch zur Variablenaggregation (Indexbildung), Ausreißer- und missing data-Behandlung, u.a. Je höher der Arbeitsaufwand für Dokumentation und Kodebuch bzw. Studienbeschreibung für die Autoren, desto geringer wird vielfach die Bereitschaft zur Archivierung sein.

Erwägenswert ist folglich, einen für die Autoren einfacheren und insgesamt weniger arbeitsaufwendigen, *zweiten Standard* der Dokumentation von Datenfiles gelten zu lassen, d.h. die Kombination von .sav und .pdf Files. Viele der Da-

² Auch die auf der eigenen [Homepage](#) archivierte Datensätze der ehemaligen Forschungsgruppe Psychophysiologie des Freiburger Psychologischen Instituts (26 Studien mit 49 Datensätzen) sind als Kombination von .sav und .pdf Files dokumentiert, und in Anlehnung an die entsprechenden Hinweise bei PsyDok werden vor dem Download nur die Adresse, die Institution und der Nutzungszweck registriert.

tengeber und der Datenehmer in der Psychologie werden wahrscheinlich diese pragmatische Variante bevorzugen, d.h. vorzugsweise SPSS-Dateien, wobei für eine überschaubare Zukunft das Interesse des SPSS-Lizenzgebers an einer Kontinuität der Lesbarkeit anzunehmen ist.

Primärdaten psychologischer Tests

Im Einvernehmen mit den Testautoren, dem Testkuratorium und den Testverlagen könnten Gesichtspunkte für die Archivierung der Primärdaten psychologischer Tests formuliert werden und ggf. Testautoren häufig verwendeter Tests für diese Konzeption gewonnen werden, insbesondere wenn die Testentwicklung durch öffentliche Gelder (Universitäten, DFG) ermöglicht wurde. Im Kern geht es um die zur Testkonstruktion verwendete Datenbasis (Itemebene), die vielfach aus einer mehr oder minder großen, repräsentativ erhobenen Stichprobe stammt oder erst sekundär aus Gelegenheitsstichproben „repräsentativ“ zusammengesetzt wurde.

- Ist über die genaue Beschreibung der Normierungsstichprobe hinaus eine Archivierung auch der – oft zahlreichen – soziodemographischen Merkmale und der psychologischen Zusatzinformationen zu verlangen?
- Kann eine ausführliche Beschreibung der eigentlichen Testkonstruktion durch Verweis auf das publizierte Test-Manual ersetzt werden?
- Welche Gründe sprechen für oder gegen die Archivierung auch der normierten Testwerte (Normen-Tabellen des Manuals)?
- Sollte die Archivierung der eigentlichen Test-Items nur mit ihrer Nummer, jedoch ohne inhaltliche Information (Aufgabe, Itemtext) dokumentiert werden, um das Testmaterial möglichst zu schützen?
- Welche typischen Missbrauchsfälle können unter realistischen Annahmen antizipiert werden?
- Was würde eine Sperrfrist von einigen Jahren bewirken können oder ist sie überflüssig, weil sie den späteren Missbrauch (welchen?) nicht ausschließt und wissenschaftlich kontraproduktiv wäre?
- Wie können in einer Übergangsphase außer den neu publizierten psychologischen Tests rückwirkend auch die Autoren der am häufigsten verwendeten Tests zur Teilnahme motiviert werden?
- Soll die Open Access Archivierung der Datenbasis als Gesichtspunkt des TBS-TK aufgenommen werden?
- Welche Empfehlungen lassen sich für die vertraglichen Regelungen zwischen Testautoren, Testverlag und ZPID formulieren?

Sind überhaupt *besondere* vertragliche Regelungen für die Primärdaten publizierter psychologischer Test erforderlich? Die vom ZPID verwendeten Vertragstexte, d.h. der Nutzungsüberlassungsvertrag (Datengebervertrag) und der Nutzungsvertrag (Datenehmervertrag), könnten in dieser allgemeinen Form oder mit wenigen speziellen Ergänzungen durchaus genügen.

Open Data – Wie werden es die Psychologinnen/Psychologen künftig halten, grundsätzlich und insbesondere bei den mit öffentlichen Mitteln (Universitäten, DFG u.a.) gewonnenen Daten: Privatisierung und Zurückhaltung *oder* Offenheit und Teilhabe?

Literaturhinweise und Links

- American Psychological Association (2001). *Ethical standards for the reporting and publishing of scientific information*. In *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: APA.
- American Psychological Association (2002). *Ethical Principles of Psychologists and Code Of Conduct*. Zugriff am 5.10.2009. Verfügbar unter <http://www.apa.org/ethics/code2002.html>
- American Psychological Association. *Data Acquisition, Management, Sharing, and Ownership*. Zugriff am 5.10.2009. Verfügbar unter <http://www.apa.org/science/data/homepage.html>
- American Psychological Association. *Supplementing Your Article With Online Material*. Zugriff am 5.10.2009. Verfügbar unter <http://www.apa.org/journals/authors/suppmaterial.html>
- Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Zugriff am 5.10.2009. Verfügbar unter <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>
- Berufsverband Deutscher Psychologinnen und Psychologen e.V. (1998/2005). *Ethische Richtlinien*. Fassung vom 16.4.2005. Zugriff am 5.10.2009. Verfügbar unter <http://www.bdp-verband.org/bdp/verband/ethik.shtml>
- Budapest Open Access Initiative. Zugriff am 5.10.2009. Verfügbar unter <http://www.soros.org/openaccess/read.shtml>
- Clark, K.E. (1971). A critical examination of the National Information System for Psychology. *American Psychologist*, 26, 325-329.
- Cockerill, M. (2003). *Data mining Open Access research*. Zugriff am 5.10.2009. Verfügbar unter <http://www.biomedcentral.com/openaccess/archive/?page=features&issue=4>
- Corti, L. (2000). *Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research – The International Picture of an Emerging Culture* [58 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(3), Art. 2, <http://nbn-resolving.de/urn:nbn:de:0114-fqs000324>.

- Council on Governmental Relations COGR (2006). Access to and retention of research data: Rights and responsibilities. Zugriff am 5.10.2009. Verfügbar unter <http://206.151.87.67/docs/CompleteDRBooklet.htm>
- Data Documentation Initiative. Zugriff am 5.10.2009. Verfügbar unter <http://www.icpsr.umich.edu/DDI/>
- Data Sharing and Replication. Gary King's Work on and Data Sharing, Replication, and Informatics. Zugriff am 5.10.2009. Verfügbar unter <http://gking.harvard.edu/replication.shtml>
- Dehnhard, I., Mittelsdorf, B. & Bäumer, T. (2006). Manual zur forschungsbegleitenden Dokumentation. Trier: ZPID PsychData. Zugriff am 5.10.2009. Verfügbar unter http://psychdata.zpid.de/downloads/psychdata_manual_0-9.pdf
- Deutsche Forschungsgemeinschaft (1998). Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“. Vorschläge zur Sicherung guter wissenschaftlicher Praxis. Zugriff am 5.10.2009. Verfügbar unter http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf
- Deutsche Forschungsgemeinschaft (Hg.). (2005). Publikationsstrategien im Wandel? Ergebnisse einer Umfrage zum Publikations- und Rezeptionsverhalten von Wissenschaftlern unter besonderer Berücksichtigung von Open Access. Zugriff am 5.10.2009. Verfügbar unter http://www.dfg.de/dfg_im_profil/zahlen_und_fakten/statistisches_berichtswesen/open_access/download/oa_ber_dt.pdf
- Deutsche Forschungsgemeinschaft. Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme. Unterausschuss für Informationsmanagement. Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten. (Januar 2009). Zugriff am 5.10.2009. Verfügbar unter http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/veroeffentlichungen/dokumentationen/download/u_a_inf_empfehlungen_200901.pdf
- Deutsche Gesellschaft für Psychologie e.V. (1998/2004). Ethische Richtlinien. Zugriff am 5.10.2009. Verfügbar unter <http://www.dgps.de/dgps/aufgaben/003.php> und <http://www.dgps.de/dgps/aufgaben/ethikr12004.pdf>
- Deutsche Initiative für Netzwerkinformationen e.V. DINI Förderung des elektronischen Publizierens im Open Access. Zugriff am 5.10.2009. Verfügbar unter <http://www.dini.de/>
- Deutsche UNESCO-Kommission (2007). Open Access. Chancen und Herausforderungen. Ein Handbuch. Bonn: Deutsche UNESCO-Kommission. Zugriff am 5.10.2009. Verfügbar unter <http://www.unesco.de/1280.html?&L=0>
- Eid, M. & Diener, E. (Eds.). (2006). *Handbook of multimethod measurement in psychology*. Washington, DC: American Psychological Association.
- Fienberg, S.E., Martin, M.E. & Straf, M.L. (Eds.). (1985). *Sharing research data*. Committee on National Statistics, National Research Council. Washington: The National Academies Press. Zugriff am 5.10.2009. Verfügbar unter http://www.nap.edu/catalog.php?record_id=2033
- Fröhlich, G. (2009). Die Wissenschaftstheorie fordert OPEN ACCESS. *Information, Wissenschaft & Praxis*, 60 (5), 253-258. Zugriff am 5.10.2009. Verfügbar unter <http://www.b-i-t-online.de/pdf/TWP2009-5.pdf>
- Gemeinsame Erklärung der Wissenschaftsorganisationen 25.3.2009. Open Access und Urheberrecht: Kein Eingriff in die Publikationsfreiheit. Zugriff am 5.10.2009. Verfügbar unter http://www.dfg.de/aktuelles_presse/download/pi_allianz_open_access.pdf
- Herb, U. (2006): Entgeltfreier Zugang zu wissenschaftlichen Informationen. Teil 1: Schöne neue Welt des Open Access, Teil 2: Journale, Impact Factor, radikale Monopole und Karrieren in telepolis, 14. und 15. September 2006. Zugriff am 5.10.2009. Verfügbar unter <http://www.heise.de/tp/r4/artikel/23/23513/1.html> und <http://www.heise.de/tp/r4/artikel/23/23531/1.html>
- Institut für Textkritik e. V. Für Publikationsfreiheit und die Wahrung der Urheberrechte Zugriff am 5.10.2009. Verfügbar unter <http://www.textkritik.de/urheberrecht/index.htm>
- Johnson, D.H. & Sabourin, M.E. (2001). Universally accessible data bases in the advancement of knowledge from psychological research. *International Journal of Psychology*, 36, 212-220.
- Kleiner, M. (2009). "Die Freiheit der Autoren ist nicht bedroht". Fragen an den DFG-Präsidenten zur Open-Access-Initiative der Allianz der Wissenschaftsorganisationen. *Forschung & Lehre* (9.10.2009). Zugriff am 5.10.2009. Verfügbar unter <http://www.forschung-und-lehre.de/wordpress/?p=2412>
- Law, M. (2005). Reduce, reuse, recycle: Issues in the secondary use of research data. *IASSIST Quarterly*, Spring. Zugriff am 5.10.2009. Verfügbar unter <http://www.iassistdata.org/publications/iq/iq29/iqvol291law.pdf>
- Linder, D.E. & Siegel, E.R. (1976). Evaluation of innovations in APA publications activities. The case of the *Journal Supplement Abstract Service*. *American Psychologist*, 31, 26-35.
- Mruck, K., Gradmann, S. & Mey, G. (2004). Open Access: Wissenschaft als Öffentliches Gut [32 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 5(2), Art. 14, Zugriff am 5.10.2009. Verfügbar unter <http://nbn-resolving.de/urn:nbn:de:0114-fqs0402141>.
- National Science Foundation-Directorate for Social, Behavioral & Economic Sciences. Data Archiving Policy. Zugriff am 5.10.2009. Verfügbar unter <http://www.nsf.gov/sbe/ses/common/archive.jsp>
- Nelson, B. (2009). Data sharing: Empty archives. Published online 9 September 2009 | *Nature* 461, 160-163. Zugriff am 5.10.2009. Verfügbar unter <http://www.nature.com/news/2009/090909/full/461160a.html>
- OECD Committee for Scientific and Technological Policy. Science, Technology and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communiqué. Zugriff am 5.10.2009. Verfügbar unter http://www.oecd.org/document/15/0,3343,en_2649_34487_25998799_1_1_1_1,00.html
- Open Access. Der freie Zugang zu wissenschaftlicher Information. Zugriff am 5.10.2009. Verfügbar unter http://open-access.net/de/oa_informationen_der_helmholtzgemeinschaft/offener_zugang_zu_daten/
- Open Access. Artikel in Wikipedia. Zugriff am 5.10.2009. Verfügbar unter http://de.wikipedia.org/wiki/Open_Access
- Open Access Plattform *Open Access - Informationsplattform* für Psychologie. Zugriff am 5.10.2009. Verfügbar unter http://open-access.net/de/oa_in_verschiedenen_faechern/psychologie/
- Pampel, H. (2008). Bericht über den Workshop „Offener Zugang zu Daten – eine Herausforderung“ im Rahmen der Open-Access-Tage 2008 am 10.10.2008 in Berlin. Zugriff am 5.10.2009. Verfügbar unter http://www.ib.hu-berlin.de/~libreas/libreas_neu/ausgabe14/017pam.htm

- Pfeiffenberger, H. (2007). Offener Zugang zu wissenschaftlichen Primärdaten. *Zeitschrift für Bibliothekswesen und Bibliographie*, 54, Nr. 4/5, 207-210. Zugriff am 5.10.2009. Verfügbar unter http://oa.helmholtz.de/fileadmin/Links/Artikel/ZfBB_Pfeiffenberger_final.pdf
- Ruusalepp, R. (2008). Infrastructure, Planning and Data Curation. A comparative study of international approaches to enabling the sharing of research data. Version 1.6. 30. November 2008. Zugriff am 5.10.2009. Verfügbar unter http://www.dcc.ac.uk/docs/publications/reports/Data_Sharing_Report.pdf
- Schweizer, K. (1989). Eine Analyse der Konzepte, Bedingungen und Zielsetzungen von Replikationen. *Archiv für Psychologie*, 141, 85-97.
- Siegel, E. R. (1972). The Journal Supplement Abstract Service: A New Concept in Psychology Communication. Paper presented at the International Communication Association Annual Convention. Atlanta, Georgia, April 19-22, 1972. Zugriff am 5.10.2009. Verfügbar unter http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/3b/1f/03.pdf
- Spindler, G. (Hg.). (2006). *Rechtliche Rahmenbedingungen von Open-Access-Publikationen*. Universität Göttingen, e-book, PDF 3,6 MB. Göttingen: Universitätsverlag.
- Testkuratorium (2007). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Föderation Deutscher Psychologinnenvereinigungen. *Psychologische Rundschau*, 58, 25-30.
- Wicherts, J.M., Borsboom, D., Kats, J. & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, 61, 726-728.
- Wolins, L. (1962). Responsibility for raw data. *American Psychologist*, 17, 657-658.

Ausgewählte Daten- und Dokumenten-Banken

- ALLF Archiv für Lebenslaufforschung. Zugriff am 5.10.2009. Verfügbar unter <http://www.lebenslaufarchiv.uni-bremen.de/index.php?>
- Directory of open access journals. Zugriff am 5.10.2009. Verfügbar unter <http://www.doaj.org/>
- Forum Qualitative Sozialforschung FQS. Zugriff am 5.10.2009. Verfügbar unter <http://www.qualitative-research.net/index.php/fqs/index>
- GESIS - Leibniz-Institut für Sozialwissenschaften. Zugriff am 5.10.2009. Verfügbar unter <http://www.gesis.org/dienstleistungen/forschungsdatenzentren/>
- PsychData. Zugriff am 5.10.2009. Verfügbar unter <http://psychdata.zpid.de/>
- PsyDok Open-Access-Server der Virtuellen Fachbibliothek Psychologie. Zugriff am 5.10.2009. Verfügbar unter <http://psydok.sulb.uni-saarland.de/>
- QUALIDATA, das Qualitative Data Archival Resource Centre des englischen Economic and Social Science Research Council (ESRC) (siehe Corti (2000)). Zugriff am 5.10.2009. Verfügbar unter <http://www.esds.ac.uk/qualidata/about/introduction.asp>
- Sozioökonomisches Panel SOEP. Zugriff am 5.10.2009. Verfügbar unter <http://www.diw.de/deutsch/soep/29004.html>
- Übersicht von Open-Access-Zeitschriften des Zentrums für Psychologische Information und Dokumentation (ZPID). Zugriff am 5.10.2009. Verfügbar unter <http://www.zpid.de/redact/category.php?cat=292>
- Virtuelle Fachbibliothek Psychologie. Zugriff am 5.10.2009. Verfügbar unter <http://fips.sulb.uni-saarland.de/port.htm>

Endnote

ⁱ *The Journal Supplement Abstract Service* JSAS sollte die Lücke zwischen den regulären Publikationen in Zeitschriften und dem sonstigen Informationsaustausch füllen, d.h. "fugitive" literature in psychology, technical reports, invited addresses, bibliographies, literature reviews usw. aufnehmen. Der seit 1971 erscheinende jährliche *Catalog of Selected Documents in Psychology of the JSAS* wurde 1975 zur offiziellen Publikation der APA. Ursprünglich hatte Siegel (1972) propagiert: ein "New Concept in Psychology Communication" und "data archiving on a voluntary basis as an integral part of the American Psychological Association's (APA) National Information System for Psychology." Siegel zählte verschiedenste Materialien auf, nannte in seiner Vorschlagsliste jedoch an erster Stelle: a. methodological techniques, b. massive data collections, c. educational materials usw. Bemerkenswert sind die letzten Positionen: n. well designed studies that are „near replications“ und o. "well designed studies with negative results" (p. 6). Tatsächlich scheint JSAS ausschließlich Texte archiviert zu haben. So enthält der *Catalog* im Jahr 1976 (Volume 6) keinen Hinweis auf Daten. Über Einzelheiten dieses Dokumentationsdienstes wurde von Linder und Siegel (1976) berichtet: u.a. Anzahl und Umfang der Texte, Anzahl der angeforderten Kopien, Argumente für und wider JSAS (siehe auch Clark, 1971). Im Jahr 1985 gab es einen Beschluss des *Council of Representatives*, die Publikation von *Psychological Documents* (formerly the Journal Supplement Abstract Service) einzustellen: "Continued low submissions, decreasing usage, and rising costs for fulfillment of paper and microfiche copies of documents were reasons given for discontinuing publication of the alternative format publication, which was begun in 1971 as an 'experimental' publication." (siehe Editorial Note, e.g. in *Journal Experimental Psychology: General*, 1985, 114, p. 197).

Heute wird ergänzendes Material in die Publikation eines Zeitschriftenartikels einbezogen, indem entweder im Text oder im Anhang ein Link zur Internet-Seite des Autors gegeben wird. Die [APA](#) hat nun diese Verknüpfung mit den *supplementary materials* geregelt (siehe auch APA Publication Guidelines, 6th edition). – Die hier gemeinten supplementary materials sind jedoch von den primary data der Publikation zu unterscheiden.

“APA can now place supplementary materials online, linked to the published article in the PsycARTICLES database. Examples of such material: audio or video clips, oversized tables, lengthy appendixes, detailed intervention protocols, supplementary data sets.

If you are interested in including such material with your article, please state your intentions in the cover letter that accompanies your submission to the Editor, and include the material with your submission. All material submitted for publication is subject to peer review.

Provided the supplementary material is accepted for online publication, the final published version of the article will include a URL noting where the supplementary materials can be found online. It is not necessary to refer to the supplementary material in the text of your manuscript, but you may do so if you wish. Once accepted, the supplemental material is considered part of the journal's archival record and cannot be altered or removed.”

Postanschrift
Prof. Dr. Jochen Fahrenberg
Waldhofstrasse 42
D-79117 Freiburg

e-mail jochen.fahrenberg@psychologie.uni-freiburg.de
homepage <http://www.jochen-fahrenberg.de>