PsychData ZPID

I. Dehnhard, G. Schui, E. Weichselgartner

ZPID - PsychData · Universität Trier · 54296 Trier · psychdata.zpid.de · psychdata@zpid.de

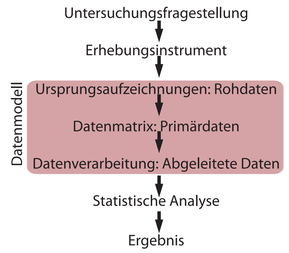
gefördert von: DFG

Einleitung

Erkenntnisgewinn in der Psychologie als empirische Wissenschaft ist seit ihren Anfängen unmittelbar an die Produktion und Verarbeitung von Daten gebunden, welche das Erleben und Verhalten von Menschen abbilden. Ungeachtet der hervorgehobenen Position des empirischen Zugangs lassen sich in der Forschungspraxis Defizite hinsichtlich des Umgangs mit Datenbeständen formulieren, die sich auf die lückenhafte Dokumentation sowie den unzureichenden Schutz vor Alterung, Verlust und Manipulation beziehen. Diese Defizite führen dazu, dass Daten selbst für diejenigen, die sie erhoben haben, innerhalb kürzester Zeit nicht mehr uneingeschränkt verfügbar sind.
Das Zentrum für Psychologische Information und Dokumentation (ZPID) bietet mit dem Psychologischen Datenarchiv PsychData ein Dienstleistungsangebot an, das Forschenden die standardisierte Dokumentation und langfristige Archivierung von Daten ermöglicht, die eindeutige Zitation dieser Daten garantiert und Informationen zu diesen Daten über eine komfortable Web-Schnittstelle bereitstellt.

Datenmodell - Archivierungsobjekt

Daten werden in der Psychologie mit unterschiedlichster Methodologie erhoben. Die Rohdaten oder Ursprungsaufzeichnungen reichen dabei von Angaben auf Fragebögen und Ratingskalen über Ton- und Videoaufzeichnungen bis hin zu umfangreichen Datenströmen aus bildgebenden Verfahren.
Um ein eindeutiges Archivierungsobjekt zu erhalten, ist eine Homogenisierung dieses heterogenen Materials notwendig. Im Forschungsprozess üblich ist dabei die Überführung der Ursprungsaufzeichnungen in Datenmatrizen, die durch Variablen einerseits, durch Fälle andererseits aufgespannt werden.
Diese als Primärdaten bezeichneten Datenmatrizen stellen das zentrale Archivierungsobjekt im Psychologischen Datenarchiv dar. Datenmatrizen haben dabei den entscheidenden Vorteil, dass sie in statistischen Analyseroutinen ohne Aufwand weiterverarbeitet werden können. In PsychData können neben den Primärdaten auch Ursprungsaufzeichnungen und abgeleitete Variablen (z.B. Skalenwerte, Indices) archiviert werden.



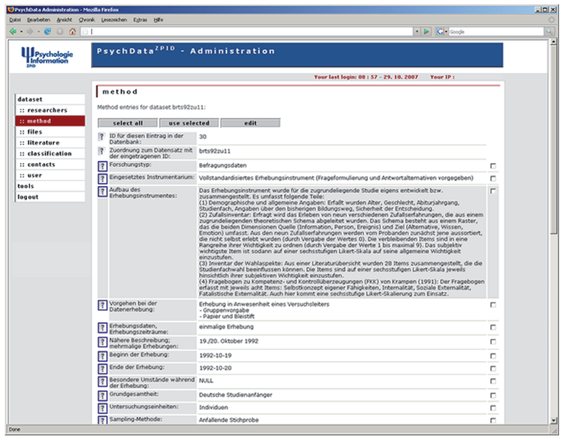
Defizite im Umgang mit Daten

Trotz der hohen Bedeutung des empirischen Zugangs in der Psychologie kann die Datenpflege – aufgrund personaler, zeitlicher oder finanzieller Ressourcenknappheit – häufig nicht optimal betrieben werden. Daraus resultiert:
- Mangelnde Rekonstruierbarkeit und Interpretierbarkeit einzelner Datenpunkte
- Mangelnde Dokumentation des Datenerhebungsprozesses
- Probleme bei der gemeinsamen Bearbeitung von Daten in größeren, insbesondere internationalen Forschergruppen
- Verlust von Datendateien aufgrund technischer Defekte oder menschlichen Versagens
- Fehlende Möglichkeit zur Publikation und Zitation teils aufwendig erhobener Datenbestände
- Keine rechtlich abgesicherte Basis zur Weitergabe von Forschungsdaten an Sekundärforscher

Dokumentation – Metadaten

Die Dokumentation der in PsychData eingestellten Datensätze beruht auf einem definierten Satz von Metadaten. Diese „Daten über Daten" umfassen die Beschreibung der Datenpunkte selbst und des Forschungskontextes und dienen insbesondere dem Retrieval der aufgenommenen Datenquellen.
Das im Projekt entwickelte Metadatenschema, das sich an international verbreiteten und akzeptierten Standards und den Besonderheiten empirisch-psychologischer Forschung orientiert, umfasst 163 Felder.

- Administrative Metadaten
- Bibliographische Metadaten
- Studienbeschreibende Metadaten
- Methodenbeschreibende Metadaten
- Filebeschreibende Metadaten
- Datenbeschreibende Metadaten



Forschungsbegleitende Aktivitäten.

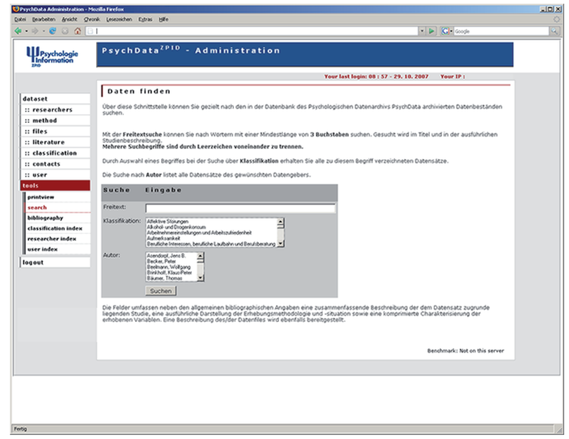
1) „Manual zur forschungsbegleitenden Dokumentation."
Ziel des Handbuchs ist es, die Forscher bei Ihrem Dokumentationsprozess zu unterstützen und so die Eingabe in ein Datenarchiv vorzubereiten. Dabei soll die Dokumentationsarbeit so weit wie möglich in den Forschungsprozess eingebunden werden, so dass möglichst wenig zusätzlicher Arbeitsaufwand entsteht. Zum Dokumentationsprozeß gehören
(1) Studiendokumentation,
(2) Variablendokumentation (Kodebuch),
(3) Dateneingabe und
(4) Archivierung.

2) "Webbasierte Tools zur Datendokumentation und -archivierung."
Maßgeschneiderte Werkzeuge sollen Forschende in die Lage versetzen, Datenbestände ökonomisch und effizient umfassend zu dokumentieren und deren langfristige Verwahrung und Zugänglichkeit sicher zu stellen. Mit Hilfe der interaktiven Oberfläche "PsychDataExtern" können Forscher Ihre Metadaten, Kodebücher und Primärdaten eingeben und verwalten.

Präsentation – psychdata.zpid.de

Während bislang ausschließlich Forschungsergebnisse in Form von Buchbeiträgen oder Zeitschriftenartikeln publiziert und damit der Fachöffentlichkeit zugänglich gemacht werden können, werden im Psychologischen Datenarchiv nunmehr die solchen Veröffentlichungen zugrunde liegenden Forschungsdaten selbst publiziert.
Die PsychData-Homepage bietet:
- Datenbankgestützte Suchoberfläche mit Freitextsuche und Suchkategorien
- Bereitstellung ausgewählter Metadaten in einer Kurz- und einer Langversion
- Bezugsmöglichkeit von Forschungsdaten und Kodebüchern via CD-ROM
- Präsentation der Daten in zitierfähiger Form

Datenbestand:
Aktuell (Herbst 2007) werden etwa 23 Millionen Datenpunkte in 55 Datensätzen aus 30 Studien vorgehalten.



Fazit – Serviceangebot von PsychData

Die Forschungskultur und Forschungsökonomie in der Psychologie profitiert im Sinne der „Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen" von einem freien Zugriff auf unterschiedlichste Informationsinhalte. Vor diesem Hintergund lässt sich als übergreifendes Ziel von PsychData formulieren:
Jeder Forschende wird in die Lage versetzt, anfallende Datenbestände ökonomisch und effizient – d.h. mit vertretbarem Aufwand – umfassend zu dokumentieren, langfristig sicher zu verwahren und auf einer informativen Oberfläche zu präsentieren.

PsychData unterstützt Forschende:
- Bei der Dokumentation eigener Datensätze – unabhängig von Forschungsgebiet und methodischem Zugang – durch Bereitstellung geprüfter Standards und Werkzeuge sowie Beratung durch unsere Mitarbeiter
- Bei der Archivierung bedeutsamer Datensätze durch Bereitstellung von Speicherplatz auf File-Servern des ZPID und eine elaborierte Datensicherungsstruktur
- Beim Aufbau einer institutsinternen Dokumentations- und Archivierungsumgebung durch Implementierung von und Einweisung in die PsychData-Administrationsumgebung
- Bei der Außendarstellung ihrer Forschungsbemühungen und Intensivierung der Rezeption der Daten durch Publikation auf der PsychData-Homepage

Workshop 2008

Auf dem „International Congress of Psychology" vom 20.-25. Juli 2008 in Berlin bieten wir zum Thema Primärdatenarchivierung in der Praxis den Workshop „Sharing Research Data: Tools and Procedures" an.