
Projekt DataWiz: Entwicklung eines Assistenzsystems zum Management psychologischer Forschungsdaten

Martin Kerwer¹, Ronny Bölter², Ina Dehnhard³, Armin Günther⁴, Erich Weichselgartner⁵

1,2,3,4,5 Leibniz-Zentrum für Psychologische Information und Dokumentation, ZPID, Trier

Zusammenfassung. Über Disziplinen hinweg werden Nullbefunde kaum veröffentlicht und Daten selten geteilt. Selbst auf Anfrage sind Daten, die Publikationen zugrunde liegen, häufig nicht mehr verfügbar. Neben den damit verbundenen wissenschaftstheoretischen Problemen führt dies zu einem ineffizienten Einsatz öffentlicher Gelder, weswegen Forschungsförderer und Fachorganisationen verstärkt für einen besseren Umgang mit Forschungsdaten und deren freie Verfügbarkeit eintreten. Die (deutsche) Psychologie nimmt hierbei mit ihren neuen Empfehlungen zum Umgang mit Forschungsdaten eine Vorreiterrolle ein. Um den Anforderungen, die sich durch ein hochwertiges Forschungsdatenmanagement stellen, gerecht werden zu können, benötigen Forschende aber auch Werkzeuge, die ihnen die Bewältigung dieser Anforderungen mit möglichst wenig Zusatzaufwand ermöglichen. Die Entwicklung und Bereitstellung solcher Werkzeuge stellt eine disziplinübergreifende Herausforderung dar. Mit dem Projekt DataWiz wird die Entwicklung eines Forschungsdatenmanagement-Werkzeugs, das diese Herausforderung in der Psychologie angeht, als Fallstudie vorgestellt.

Das derzeit am ZPID entwickelte computergestützte Assistenzsystem DataWiz unterstützt Psychologen/innen in ihrem täglichen Forschungsdatenmanagement wissensbasiert und prozedural. Wissensbasierte Unterstützung findet innerhalb der Anwendung in Form einer durchgängigen, kontextsensitiven Bereitstellung von Informationen zu forschungsdatenrelevanten Themenbereichen (Datenschutz, Urheberrecht, Qualitätssicherung, usw.) statt.

Gleichzeitig wird Forschenden mit DataWiz aber auch ein Werkzeug zur Verfügung gestellt, das ihnen bei der Bewältigung typischer Aufgaben des Forschungsdatenmanagements assistiert. Die Webanwendung verbindet grundlegende Funktionen, die das kollaborative Arbeiten in einer virtuellen Forschungsumgebung ermöglichen, wie einer abgestuften Nutzerrechtevergabe und ein entsprechend abgestufter Zugang zu abgelegten Dateien, mit prozeduralen Funktionen, die Nutzer/innen im Forschungsprozess begleiten. Dazu zählen eine angeleitete Planung des Forschungsdatenmanagements, die fachspezifische Dokumentation der Datenerhebung und der erhobenen Forschungsdaten, sowie die Aufbereitung der Forschungsdaten und ihre Übergabe an fachlich fundierte und etablierte Datenarchive. DataWiz zeichnet sich hierbei unter anderem durch den einfachen Import von Datenmatrizen und zugehöriger Metadaten aus disziplinüblicher Software, einer Verknüpfung von Datenmanagementplanung, Codebucherstellung und Datenerhebungsdokumentation, sowie die Unterstützung des Forschenden in der Datenaufbereitung und Versionierung aus.

Neben einer frei zugänglichen, am ZPID gehosteten Version, wird Nutzer/innen die Möglichkeit gegeben, eigene Versionen der Anwendung an ihrer Institution zu betreiben. Um auch das interdisziplinäre Nachnutzungspotenzial der Daten zu erhöhen, werden disziplinübergreifende Metadatenstandards unterstützt. Die Übergabe der Daten an Forschungsdatenarchive verwandter Disziplinen wird damit stark vereinfacht. Darüber hinaus wird der Quellcode des Programms der Community offen zur Verfügung gestellt, wodurch DataWiz eine potentielle Basis fächerübergreifender Entwicklungen darstellt.

Schlagwörter. Forschungsdaten, Data Sharing, Data Management

Theoretischer Hintergrund

Forderung nach Data Sharing

Die Forschung der letzten Jahre hat wiederholt gezeigt, dass die Replizierbarkeit wissenschaftlicher Befunde zu wünschen übriglässt (z.B. Open Science Collaboration, 2015). Eine Ursache hierfür ist der als File-Drawer-Problem bekannte Sachverhalt, dass Nullbefunde häufig nicht berichtet werden. Dieser führt zu Publication Bias - einer Verzerrung der in der Literatur berichteten Ergebnisse von Studien. Diese Verzerrung ist aus methodischer Sicht höchst problematisch, da die Ergebnisse von Sekundäranalysen nicht zuverlässig für Publication Bias korrigiert werden können (Scargle, 2000). Ein weiteres Problem liegt darin, dass häufig auch nicht die kompletten Ergebnisse einer Studie berichtet werden. Stattdessen findet vermehrt eine auszugsweise Berichterstattung signifikanter Ergebnisse statt (sog. P-Hacking, Simonsohn, Nelson, Simmons, 2014). Diese verzerrte Berichterstattung innerhalb von Studien wird auch als Reporting Bias bezeichnet.

Beide Problemstellungen können dadurch gelöst werden, dass die Hypothesen einer Studie präregistriert und die kompletten Rohdaten, die der Publikation zugrunde liegen, veröffentlicht werden. Abweichungen vom ursprünglichen Analyseplan werden damit transparent gemacht und alternative Spezifikationen der Analysen können auf Grundlage dieser Daten getestet werden. Das Potenzial der Bereitstellung von Forschungsdaten – des Data Sharing – reicht aber weit über diesen Anwendungshorizont hinaus, wenn Daten nicht nur zur Qualitätssicherung verfügbar gemacht werden, sondern auch zur weiteren Nutzung bereitgestellt werden. Die Vorteile des Data Sharing umfassen in diesem Fall unter anderem eine verbesserte Nutzung der Ressourcen, die für Forschungsvorhaben aufgebracht werden, die Vermeidung einer unnötigen Beanspruchung von Versuchsteilnehmern und die Erhaltung einmaliger, nicht replizierbarer Datenbestände.

Die breite Wahrnehmung der Replikationskrise und des potenziellen Nutzens des Data Sharing hat dazu geführt, dass die freie Verfügbarkeit von Forschungsdaten in den letzten Jahren immer stärker ins Blickfeld wissenschaftlicher Akteure gerückt ist. Zu diesen Akteuren zählen unter anderem nationale und internationale Förderorganisationen, wie die Deutsche Forschungsgemeinschaft (Richtlinien zum Umgang mit Forschungsdaten, Deutsche Forschungsgemeinschaft, 2015) und die Europäische Union (Ausweitung des Open Data Pilot des Horizon 2020 Programms, European Commission, 2016, Art 29.3.), Publisher (z.B. PLOS¹, Society for Judgment and Decision Making²) und wissenschaftliche Fachgesellschaften. Die Deutsche Gesellschaft für Psychologie (DGPs) beispielsweise verabschiedete fachspezifische Empfehlungen zum Datenmanagement in der Psychologie (Schönbrodt, Gollwitzer, & Abele-Brehm, 2016), die sich für eine freie Verfügbarkeit der Daten, die Publikationen zugrunde liegen, einsetzen. Vor dem Hintergrund, dass die Bereitschaft Daten zu teilen in der Psychologie unverändert niedrig ist (Dehnhard, Weichselgartner und Krampen, 2013, Vanpaemel, Vermogen, Deriemaecker und Storms, 2015, Wicherts, Borsboom, Kats und Molenaar, 2006), sollte die Bedeutung dieser Empfehlungen nicht unterschätzt werden.

1 <http://journals.plos.org/plosone/s/data-availability> (abgerufen am 08.02.17)

2 <http://journal.sjdm.org/>(abgerufen am 08.02.17)

Forderung nach Datenmanagement

Zunehmend wird außerdem erkannt, dass die reine technische Verfügbarkeit von Forschungsdaten nicht genügt, um ihre Nachnutzbarkeit zu ermöglichen. Neben der technischen Verfügbarkeit der Daten muss nämlich hierzu auch die langfristige Verständlichkeit der Daten sichergestellt werden. Für die Erhaltung der Interpretierbarkeit bedarf es zusätzlicher Informationen zu den Daten, sogenannter Metadaten. Diese können zum Teil genereller und damit fächerübergreifender Natur sein, in den meisten Disziplinen werden aber auch fachspezifische Informationen benötigt. Die Forderung nach einem fachgerechten Datenmanagement wird deswegen immer lauter (z.B. Council of the European Science, 2016, Schönbrodt, Gollwitzer, & Abele-Brehm, 2016). Fachspezifische Besonderheiten in der Psychologie, die spezifisches Wissen und Kompetenzen für das Datenmanagement verlangen, sind zum Beispiel der Einsatz einer großen Bandbreite an (häufig kaum dokumentierten) Datenerhebungsverfahren, ein hoher Anteil an Daten aus Studien ohne institutionalisiertes Forschungsdatenmanagement, oder das Vorliegen personenbezogener Daten, die besondere datenschutzrechtliche Kenntnisse erfordern (Weichselgartner, 2011). Das entsprechende Wissen muss Forschenden so einfach wie möglich zur Verfügung gestellt werden. Fraglich ist aber auch, ob Forscher/innen ohne prozessorientierte Unterstützung in der Lage sind, datenmanagementbezogene Aufgaben fachgerecht zu bewältigen. So konnten Vines et al. (2014) zeigen, dass die Verfügbarkeit von Daten, die Publikationen zugrunde liegen, mit zunehmendem Alter der Publikation stark abnimmt. Dies galt unabhängig davon, ob Forscher gewillt waren ihre Daten zu teilen oder nicht. Die Kompetenz von Forschern, ihre Forschungsdaten (auch für die eigene Nutzung) zu erhalten, muss damit angezweifelt werden.

Wir müssen uns also fragen, wie wir Forschende dazu befähigen können den Anforderungen, die ein fachgerechtes Datenmanagement an sie stellt, gerecht zu werden. Dabei ist der (wahrgenommene) Arbeitsaufwand, der hierfür notwendig ist, eine kritische Größe. Denn auch wenn Data Sharing mittel- und langfristig gesehen helfen kann, Ressourcen einzusparen, bindet ein sachgemäßes Forschungsdatenmanagement zunächst einmal Ressourcen. Die Fragmentierung datenmanagementbezogener Informationen und unterstützender Angebote (z.B. zur Planung des Forschungsdatenmanagements, zur Dokumentation und Aufbereitung der Daten selbst oder zur Übergabe der Daten an ein Archiv) trägt sicher zum wahrgenommenen Arbeitsaufwand bei. Eine Reduktion dieses Aufwands kann dadurch erreicht werden, dass Datenmanagement frühzeitig in den Forschungsprozess integriert wird (Gutmann et. al., 2009). Die frühzeitige Integration in den Forschungsprozess erlaubt es, Informationen dort abzugreifen, wo sie entstehen. Auf diese Weise wird die mühsame Rekonstruktion von Informationen am Ende eines Forschungsprojekts vermieden.

Darüber hinaus muss berücksichtigt werden, dass Forscher ein berechtigtes Interesse daran haben, zunächst selbst in ihrer Arbeitsgruppe die Kontrolle über ihre Forschungsdaten zu behalten. Diese Rückmeldung aus der Fachgemeinschaft hat das ZPID im Betrieb des Forschungsdatenzentrums PsychData immer wieder erhalten und auch die Teilnehmer des DataWiz-KickOff-Workshops haben sich entsprechend geäußert. Wichtig ist deshalb, dass Forschende die Bedingungen des Datenzugangs selbst festlegen können (Simukovic, Kindling und Schirnbacher, 2013).

Anforderungen an Forschungsdatenmanagement-Tools

Der Druck auf Forschende, hochwertige Daten bereitzustellen, muss also von Angeboten begleitet werden, die Forschende bei der Bewältigung dieser neuen Anforderungen unterstützen. Forschungsdatenmanagement-Werkzeuge können dies leisten, wenn sie empirisch arbeitende Forscher und Forscherinnen in die Lage versetzen, selbstständig ein hochwertiges Forschungsdatenmanagement durchzuführen. Folgenden Anforderungen an die Entwicklung solcher „Assistenzsysteme“ lassen sich aus unseren vorangegangenen Überlegungen ableiten:

- Forschungsdatenmanagement-Werkzeuge müssen nicht nur dabei helfen, die Verfügbarkeit der Daten, sondern auch ihre Interpretierbarkeit sicherzustellen.
- Fachspezifische Erfordernisse müssen berücksichtigt werden.
- Forschende brauchen einen einfachen Zugriff (single-point access) auf relevante inhaltliche Informationen zum Datenmanagement (wissensbasierte Unterstützung).
- Forschende brauchen Unterstützung bei der Bewältigung zentraler Aufgaben des Datenmanagements (prozedurale Unterstützung).
- Forschende müssen frühzeitig und über den gesamten Forschungsprozess hinweg unterstützt werden.
- Forschende müssen die Kontrolle über ihre Daten (zunächst) behalten. Die Freigabe der Daten für andere (Data Sharing) darf nur optional sein.
- Insgesamt ist der Aufwand, den Forschende in ein adäquates Forschungsdatenmanagement investieren müssen, zu minimieren.

Aus Perspektive der Forschungsdateninfrastruktur (insbesondere Forschungsdatenzentren, Repositorien) hätte ein durch entsprechende Tools assistiertes Forschungsdatenmanagement gegenüber einem nicht-assistierten Datenmanagement den wesentlichen Vorteil, dass der erforderliche Archivierungsaufwand bei der Übernahme der Daten durch eine höhere Qualität der Einreichungen reduziert werden würde.

Das Assistenzsystem DataWiz

Am Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) wird im DFG-geförderten Projekt DataWiz derzeit ein gleichnamiges webbasiertes Assistenzsystem entwickelt, das sich an diesen Anforderungen orientiert und zum Ziel hat, Psychologinnen und Psychologen zur selbstständigen Umsetzung eines fundierten Forschungsdatenmanagements im Forschungsprozess zu befähigen. Im Folgenden werden wir auf den Programmaufbau, technologische Grundlagen, sowie Nachnutzungspotenziale und Perspektiven, die sich aus der Entwicklung ergeben, eingehen.

Programmaufbau

DataWiz gliedert sich in ein System zur wissensbasierten Unterstützung, ein System zur prozeduralen Unterstützung und übergreifende Systemfunktionen, die das kollaborative Arbeiten im Team ermöglichen. Die prozedurale Unterstützung entlang des Forschungsdatenlebenszyklus wird

so implementiert, dass Redundanzen (Mehrfacheingabe gleicher Informationen) möglichst komplett vermieden werden, während gleichzeitig die wissensbasierte Unterstützung kontextsensitiv in die prozedurale Unterstützung integriert wird.

Wissensbasierte Unterstützung

Zur inhaltlichen Unterstützung der Forschenden wurde eine auf die Bedürfnisse der Psychologie ausgerichtete Wissensbasis zum Datenmanagement erstellt, die den Nutzern grundlegende Informationen zu forschungsdatenrelevanten Themenbereichen (Datenschutz, Urheberrecht, Qualitätssicherung, usw.) in der Psychologie bereitstellt und als kontextsensitives Hilfesystem fungiert. Dies führt zu einer erheblichen Aufwandsreduktion für Forschende, da die Zusammenstellung und Bewertung dieser Informationen sonst erhebliche Ressourcen erfordert. Außerdem werden Fehlentscheidungen vermieden, die das spätere Data Sharing möglicherweise erschweren oder sogar verhindern (z.B. aufgrund der Verwendung ungeeigneter, überrestriktiver Einwilligungserklärungen bei der Erhebung psychologischer Daten).

Prozessorientierte Unterstützung

Die prozessorientierte Unterstützung der Forschenden bei der Bewältigung grundlegender Aufgaben des Forschungsdatenmanagements beinhaltet Funktionalitäten zur Planung des Forschungsdatenmanagements, Datenablage, Datendokumentation, Qualitätskontrolle und zum Data Sharing sowie die Integration dieser Funktionalitäten in ein Gesamtsystem.

Datenmanagementplanung

Zu Beginn ihres Projekts unterstützt DataWiz Forscher und Forscherinnen in der Planung ihres Forschungsdatenmanagements. Informationen der Wissensbasis fungieren hierbei als Hilfestellung, während die in der Planungsphase abgefragten Informationen sich an den Auflagen großer Förderorganisation (DFG, EU, BMBF) orientieren.

Datenablage

Nach der Datenerhebung können Forschende ihre Forschungsdaten in DataWiz ablegen. DataWiz unterstützt die herstellerunabhängige Verwaltung von Forschungsdaten, die als rechteckige Datenmatrizen vorliegen, in offenen Formaten. Um den Aufwand des Imports und Exports von Datenmatrizen zu reduzieren und die Integration in den psychologischen Forschungsprozess voranzutreiben, wurden Import- und Exportroutinen für die psychologieübliche Software SPSS sowie Dateien mit festem Trennzeichen (delimiter-separated values) entwickelt.

Datendokumentation

Ein Codebucheditor erlaubt die Dokumentation der hochgeladenen Forschungsdaten durch die Erfassung und Bearbeitung von Metadaten auf Variablenebene. Importroutinen aus disziplinspezifischer Software vereinfachen diesen Prozess erheblich.

Die Studiendokumentation stellt sicher, dass der Kontext der Datenerhebung verstanden wird und Daten damit bewertet oder nachgenutzt werden können. In der Entwicklung wurde hierbei auf eine größtmögliche Kompatibilität zum Metadatenschema des durch das ZPID-betriebenen Forschungsdatenzentrums PsychData und einschlägigen fachspezifischen Berichtsstandards geachtet. Interoperabilität zu verwandten Disziplinen und Maschinenlesbarkeit der Metadaten wird durch Kodierung nach dem Data Documentation Initiative (DDI)-Metadatenstandard³ gefördert.

Qualitätskontrolle

Zur Unterstützung der Qualitätssicherung werden bei Eingabe von Metadaten in Codebücher automatisiert Konsistenzprüfungen durchgeführt, wodurch Inkonsistenzen vermieden werden. Gleichzeitig sorgt ein automatisiertes System zur Versionierung von Forschungsdaten und Codebüchern dafür, dass die Entstehung des Datensatzes nachvollzogen werden kann.

Data Sharing

Die Weitergabe der Daten und Metadaten im Forschungsprojekt wird durch die übergreifenden Systemfunktionen zum kollaborativen Arbeiten unterstützt. Exportroutinen zur Datenübergabe an das psychologische Forschungsdatenzentrum PsychData sind bereits im Design des Assistenzsystems angelegt, da das Metadatenschema des Assistenzsystems auf einem Ausbau des PsychData-Metadatenschemas basiert. Für die Übergabe von Forschungsdaten und Metadaten aus DataWiz an andere Archive werden Exportfunktionen entwickelt, die Daten und Metadaten in offenen Formaten zusammenstellen und den Arbeitsaufwand für Forscher damit reduzieren.

Integration des Workflows

Der Arbeitsaufwand für Forschende wird durch eine Integration des Forschungsdatenmanagement-Workflows weiter reduziert, da keine Redundanzen zwischen den verschiedenen in DataWiz angelegten prozeduralen Funktionalitäten anfallen. So können Nutzer beispielsweise in der Planung ihres Forschungsdatenmanagements angeben, welche Konstrukte (nicht direkt beobachtbare Personenmerkmale) in ihrem Projekt untersucht werden sollen. Im weiteren Forschungsprozess kann dann ausgewählt werden, in welcher Teil-Untersuchung welches Konstrukt erhoben wurde. Schließlich wird bei der Erstellung eines Codebuchs dokumentiert, welche Variable eines hochgeladenen Datensatzes sich welchem Konstrukt zuordnen lässt. Gleichzeitig erhält der Forschende über die Wissensbasis des Programm Informationen dazu, welche Vokabulare und Ontologie ihm zur Verfügung stehen, um das Konstrukt in einer nachnutzbaren Form zu beschreiben.

3 <http://www.ddialliance.org/> (abgerufen am 08.02.17)

Übergreifende Systemfunktionen

Natürlich fallen im Forschungsprozess neben den eigentlichen Forschungsdaten Projektoutputs an, die für das Verständnis von Daten oder Analysen bzw. für die Zusammenarbeit im Projekt notwendig sind. DataWiz verfügt deshalb über ein System zur Verwaltung beliebiger digitaler Zusatzmaterialien. Um als kollaborative Arbeitsplattform nutzbar zu sein, wurde zudem eine abgestufte Zugriffs- und Rechtekontrolle implementiert, die die Vergabe differenzierter Zugriffsrechte auf Forschungsdaten, Metadaten und Dateien an andere Nutzer erlaubt.

Technologie

Die Implementierung von DataWiz erfolgt in Java (Enterprise Edition) unter Hinzunahme von weiteren, gängigen Technologien. Hierzu zählt im Frontend Bereich neben den üblichen Standards wie HTML5, CSS3 und JavaScript vor allem die Hinzunahme von Bootstrap, jQuery und einigen kleineren jQuery Libraries. Bei der Entwicklung des Backend Bereiches wurde neben Java/Java EE vor allem weite Teile des Spring MVC Frameworks verwendet. Zusätzlich kommt die Spring Security Erweiterung zum Einsatz, um das System vor Angriffen abzusichern. Zur Speicherung hochgeladener Datensätze in Dateiform (und anderer Zusatzmaterialien) wurde auf das Minio Cloud Storage System zurückgegriffen. DataWiz legt nicht nur die Originaldatensätze in diesem System ab, sondern führt während des Importprozesses auch Konsistenzchecks auf den importierten Datenmatrizen und Codebüchern aus, die anschließend in einer MySQL Datenbank abgelegt werden. Diese redundante Speicherung hat den Vorteil, dass die Originaldatei erhalten bleibt und dem Datengeber kontinuierlich zur Verfügung steht. Gleichzeitig liegen die Daten in einer aufbereiteten Form vor, welche für den späteren Export in Langzeitarchive genutzt werden kann. Durch dieses Vorgehen ist es möglich, Änderungen an den Datensätzen von DataWiz nachzuvollziehen (Data Provenance). Die Wissensbasis wird im Content-Management-System WordPress entwickelt. Prozedurale Funktionalitäten sind auf Deutsch und Englisch implementiert, während die Wissensbasis in der internationalen Psychologie-Fachsprache Englisch bereitgestellt wird.

Entwicklungsstand und Release

Ab Anfang 2017 ist eine erste Fassung der Wissensbasis unter www.datawiz.de zugänglich. Im Laufe des Frühjahrs 2017 sollen Import- und Export-Prozeduren für SPSS-Dateien als eigenständige Komponente auf GitHub bereitgestellt werden. Das Release der finalen Version und die freie Bereitstellung des Source Codes ist für Ende 2017 geplant.

Resümee

Der psychologischen Community wird DataWiz am Projektende durch eine frei zugängliche am ZPID gehostete Version zur Verfügung gestellt. DataWiz bietet aber darüber hinaus erhebliches Nachnutzungspotenzial gerade im Hinblick auf die zukünftige Entwicklung und Bereitstellung von Forschungsdatenmanagement-Werkzeugen im interdisziplinären Kontext.

Nachnutzungspotenziale

Durch die externe Bereitstellung des Source Code ist das DataWiz System zur Nachnutzung zugänglich. Nachnutzungspotenzial besteht hierbei (1) im Betrieb lokaler Installationen von DataWiz an der eigenen Institution, (2) in der Adaption von DataWiz (als prototypische Lösung) an andere Disziplinen und der Anpassung des Programms gemäß den Besonderheiten dieser Disziplinen und (3) in der Nachnutzung bestimmter Module des Source Codes, wie den Import- und Exportfunktionalitäten von SPSS-Dateien, unabhängig vom Gesamtkontext des Projekts.

Des Weiteren wird das interdisziplinäre Nachnutzungspotenzial der durch Psychologen in DataWiz erzeugten Forschungsdaten gestärkt, da aus DataWiz exportierte Metadaten und Daten in nicht-proprietären Formaten zugänglich sind. Die Interoperabilität der in DataWiz dokumentierten Daten wird darüber hinaus durch Unterstützung des sozialwissenschaftlichen DDI-Metadatenstandards erheblich verbessert.

Literaturangaben

Council of the European Science. 2016. *Council conclusions on the transition towards an open science system*. Online verfügbar unter <http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/en/pdf>. Zuletzt geprüft am 08.02.2017.

Dehnhard, I., E. Weichselgartner, und G. Krampen. 2013. "Researcher's willingness to submit data for data sharing: A case study on a data archive for psychology". *Data Science Journal* 12:172-180. doi: 10.2481/dsj.12-037.

Deutsche Forschungsgemeinschaft. 2015. *Leitlinien zum Umgang mit Forschungsdaten*. Online verfügbar unter http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf. Zuletzt geprüft am 08.02.2017.

European Commission. 2016. H2020 Programme: AGA – annotated model grant agreement (Version 2.2). Online verfügbar unter http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf. Zuletzt geprüft am 08.02.2017.

Gutmann, M. P., M. Abrahamson, M. O. Adams, M. Altman, C. Arms, K. Bollen, ... C. H. Young 2009. "From preserving the past to preserving the future: The Data-PASS project and the challenges of preserving digital social science data". *Library Trends* 57 (3): 315-337. doi:10.1353/lib.0.0039

Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science* 349 (6251): aac4716-1 - aac4716-8. <http://doi.org/10.1126/science.aac4716>.

Scargle, J. D. 2000. "Publication bias: The "File Drawer" problem in scientific inference". *Journal of Scientific Exploration* 14: 91-106.

- Schönbrodt, F., M. Gollwitzer, und A. Abele-Brehm. 2016. *Data management in psychological science: Specification of the DFG guidelines*. Online verfügbar unter http://www.dgps.de/fileadmin/documents/Empfehlungen/Data_Management_in_Psychological_Science_20160928.pdf. Zuletzt geprüft am 08.02.2017.
- Simonsohn, U., L. D. Nelson, und J. P. Simmons. 2014. „P-curve: a key to the file-drawer”. *Journal of Experimental Psychology: General* 143 (2): 534.
- Simukovic, E., M. Kindling, und P. Schirmbacher. 2013. *Forschungsdaten an der Humboldt-Universität zu Berlin: Bericht über die Ergebnisse der Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin* (Umfragebericht, Version 1.0). Online verfügbar unter <http://nbn-resolving.de/urn:nbn:de:kobv:11-100213001>. Zuletzt geprüft am 08.02.2017.
- Vanpaemel, W., M. Vermorgen, L. Deriemaecker, und G. Storms. 2015. „Are we wasting a good crisis? The availability of psychological research data after the storm”. *Collabra* 1 (1): 1–5. <http://doi.org/10.1525/collabra.13>.
- Vines, T. H., A. Y. K. Albert, R. L. Andrew, F. Débarre, D. G. Bock, M. T. Franklin, ... D. J. Rennison. 2014. The availability of research data declines rapidly with article age. *Current Biology* 24(1): 94–97.
- Weichselgartner, E. 2011. *Disziplinspezifische Aspekte des Archivierens von Forschungsdaten am Beispiel der Psychologie (RatSWD Working Paper Series, Nr. 179)*. Berlin: Rat für Sozial- und Wirtschaftsdaten.
- Wicherts, J. M., D. Borsboom, J. Kats, und D. Molenaar. 2006. „The poor availability of psychological research data for reanalysis”. *The American Psychologist* 61 (7): 726-728. doi:10.1037/0003-066X.61.7.726