Handbuch Forschungsdatenmanagement

Herausgegeben von
Stephan Büttner, Hans-Christoph Hobohm, Lars Müller

BOCK + HERCHEN Verlag
Bad Honnef
2011

Die Inhalte dieses Buches stehen auch als Online-Version zur Verfügung:
www.forschungsdatenmanagement.de
Die Onlineversion steht unter folgender Creative-Common-Lizenz:
„Attribution-NonCommercial-ShareAlike 3.0 Unported“
http://creativecommons.org/licenses/by-nc-sa/3.0/

ISBN 978-3-88347-283-6

BOCK+HERCHEN Verlag, Bad Honnef
Printed in Germany

3.2 Strategien bei der Veröffentlichung von Forschungsdaten

Sünje Dallmeier-Tiessen
CERN / Humboldt Universität zu Berlin

Zusammenfassung

Forschungsdaten liegen in Abhängigkeit der Disziplinen in vielfältigen Formen und Formaten vor. Sie sind in allen Disziplinen Teil des wissenschaftlichen Erkenntnisprozesses. Als digitales Informationsobjekt sind sie komplex und bislang wenig studiert. Mit den Möglichkeiten neuer Informationstechnologien wurden in den letzten Jahren neue Wege in der Publikation von Forschungsdaten beschritten. Mit Fokus auf die Naturwissenschaften werden im Folgenden drei Publikationsmodelle beschrieben: Die Veröffentlichung von Forschungsdaten als eigenständiges Objekt in einem Forschungsdatenrepositorium, die Veröffentlichung von Forschungsdaten mit textueller Dokumentation und die Veröffentlichung von Forschungsdaten als Anreicherung einer interpretativen Text-Publikation.

3.2.1 Einführung

Das Verständnis des Begriffes Forschungsdaten variiert je nach Disziplin. Möchte man sich dem Begriff disziplinübergreifend annähern, so lässt sich dies z. B. über die Positionierung der Forschungsdaten im wissenschaftlichen Arbeitsprozess versuchen. Forschungsdaten sind der zentrale Gegenstand des wissenschaftlichen Erkenntnisprozesses. Ausgehend von bereits veröffentlichten Ergebnissen der Community werden sie ausgewertet und eingeordnet. Im Rahmen des Publikationsprozesses werden sie anschließend beschrieben. Diese Beschreibung in textueller Form stellt traditionell eine Interpretation und Diskussion der Daten dar. Dabei sind je nach Disziplin unterschiedliche Kulturen etabliert, die auf den heterogenen Formen und Formaten der Daten basieren. Während in den Naturwissenschaften beispielsweise oftmals ein Untersuchungsgegenstand durch unterschiedliche Methoden erhoben und ausgewertet wird, wird in den Geisteswissenschaften häufig auf bereits publizierte Untersuchungsgegenstände wie historische Text- und Bildmaterialien Bezug genommen, die traditionell unter dem Begriff „Quellenmaterial“ gefasst werden.

Am Beispiel der Biologie wird die Heterogenität auch innerhalb der Disziplinen deutlich: In dieser Disziplin können z. B. numerische Daten von Laborexperimenten, audiovisuelle Objekte wie Fotos oder Videosequenzen von Tierbeobachtungen, komplexe Modelle von Simulationen (z. B. von Umweltveränderungen) als Forschungsdaten bezeichnet werden.

Da eine detaillierte interdisziplinäre Definition des Begriffs Forschungsdaten schwerfällt, werden für die folgende konzeptionelle Betrachtung der Publikation von wissenschaftlichen Daten, Forschungsdaten als digitale Informationsobjekte betrachtet, die während des Forschungsprozesses entstehen und als Grundlage einer interpretativen wissenschaftlichen Textpublikation dienen.

Mit der rasanten Entwicklung der Informationstechnologien haben sich in den letzten Jahren neue Wege im Umgang mit wissenschaftlichen Daten eröffnet, die nach dem Verständnis von Jim Gray zu einem neuen Paradigma des wissenschaftlichen Arbeitens geführt haben:

„The world of science has changed, and there is no question about this. The new model is for the data to be captured by instruments or generated by simulations before being processed by software and for the resulting information or knowledge to be stored in computers. Scientists only get to look at their data fairly late in this pipeline. The techniques and technologies for such data-intensive science are so different that it is worth distinguishing data-intensive science from computational science as a new, *fourth paradigm* for scientific exploration.“ (Hey et al., 2009, S. xix) Mit diesem „*fourth paradigm*“ entstehen neue Wege im Umgang mit wissenschaftlichen Daten.

3.2.2 Diskussion um den zeitgemäßen Umgang mit Forschungsdaten

In Deutschland wird der Umgang mit Forschungsdaten vornehmlich in Bezug auf die Nachprüfbarkeit der Daten geregelt. Ein zentraler Punkt sind dabei die von der Deutschen Forschungsgemeinschaft (DFG) 1998 veröffentlichten „Vorschläge zur Sicherung guter wissenschaftlicher Praxis“. Diese sehen vor, dass „Primärdaten als Grundlagen für Veröffentlichungen […] auf haltbaren und gesicherten Trägern in der Institution, wo sie entstanden sind, für zehn Jahre aufbewahrt werden [sollen].“ (DFG, 1998). Der Aspekt der Nachnutzung gewinnt aktuell an Relevanz. 2010 hat die DFG diesen Aspekt im „Leitfaden für Antragsteller“ verankert. In diesem heißt es: „Wenn aus Projektmitteln systematisch (Mess-)Daten erhoben werden, die für die Nachnutzung geeignet sind, legen Sie bitte dar, welche Maßnahmen ergriffen wurden bzw. während der Laufzeit des Projektes getroffen werden, um die Daten nachhaltig zu sichern und ggf. für eine erneute Nutzung bereit zu stellen.“ (DFG, 2010) Andere Förderer sind in ihren Anforderungen noch konkreter, so erwartet die *National Science Foundation* (NSF) einen Datenmanagementplan von Antragsstellern (NSF, 2011).

Auf internationaler Ebene wurden in den letzten Jahren zahlreiche Empfehlungen und Vorgaben zum Umgang mit Forschungsdaten verabschiedet. Zentral sind hier die 2007 von der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD, 2002) publizierten „*Principles and Guidelines for Access*

to Research Data from Public Funding“. In den Disziplinen variiert die Bereitschaft des „*data sharings*“ deutlich. In der Genomforschung wurden 1996 die „*Bermuda Principles*“ im Rahmen des Humangenomprojektes formuliert (z. B. Marshall, 2001), nach denen Gensequenzen frei zugänglich gemacht werden sollen. Auch Fachgesellschaften wie z. B. die *American Geophysical Union* (AGU, 1993) haben Richtlinien zum Umgang mit Forschungsdaten formuliert. Die AGU-Richtlinie bezieht sich auf Daten, die Grundlage einer Publikation sind, die in einem Organ der Fachgesellschaft veröffentlicht werden. Darüber hinaus haben auch Journale, wie z. B. PLoS ONE, in ihren Veröffentlichungsrichtlinien Aussagen zu dem Thema formuliert. PLoS ONE fordert unter dem Punkt „*Sharing of Materials, Methods, and Data*“ die Zugänglichkeit der Daten, die Basis einer Publikation sind (PLoS ONE, o. J.).

3.2.3 Publikationsmodelle

Auf Basis der beschriebenen Diskussion und der Annäherung an den Begriff „Forschungsdaten“ können diese Daten als zentraler Baustein des wissenschaftlichen Diskurses betrachtet werden.

Ausgehend von der Zugänglichkeit der Daten können unterschiedliche Aspekte einer wissenschaftlichen Nachnutzung herausgestellt werden, die auch auf die Publikationsstrategie der Daten wirken. Dabei sind die folgenden Ansätze zu unterscheiden:

- Die Publikation von Daten um eine Nachprüfbarkeit der Daten zu ermöglichen.
- Die Publikation von Daten um eine Nachnutzung der Daten zu ermöglichen.

Der häufig anzutreffende Begriff des „data sharing“ spezifiziert, dass eine Publikation von Forschungsdaten zum Zwecke der Nachnutzung durch Dritte erfolgt. Nach Borgman (2010) sprechen folgende Gründe für das „*sharing*“ wissenschaftlicher Daten: „to make the results of publicly funded data available to the public, to enable others to ask questions of extant data, to advance the state of science and to reproduce research“. Die beiden erstgenannten Argumente sind aus der Perspektive der Nachnutzung zu betrachten, deren Anspruch durch öffentliche Interessen gesteuert ist. Die beiden letztgenannten Argumente sind aus Sicht der Datenproduzenten interessant und greifen das Interesse der jeweiligen wissenschaftlichen Community auf (nach Borgman, 2010).

Zentraler Akteur bei der Publikation von Forschungsdaten ist der Wissenschaftler oder das Team, das die Daten erhebt und interpretiert.

Wird eine Veröffentlichung der Daten angestrebt, so bedarf es häufig einer erweiterten textuellen Aufbereitung der Daten von Seiten der Wissenschaftler, die im Folgenden unter dem Begriff „Dokumentation“ gefasst wird. In der

Dokumentation werden die Kontextinformationen bereitgestellt, um Personen, die an der ursprünglichen Datenproduktion nicht beteiligt waren, eine Nachnutzung der Daten zu ermöglichen.

Da es in vielen Disziplinen bisher keine etablierte Kultur der Veröffentlichung wissenschaftlicher Daten gibt, ist der Aufwand einer nachnutzbaren oder nachprüfbaren Veröffentlichung der Daten eine aufwendige und zusätzliche Tätigkeit, die über den „normalen“ Wissenschaftsalltag hinausgeht. Bei der Publikation der Daten müssen z. B., je nach Disziplin, rechtliche und ethnische Rahmenbedingungen berücksichtigt werden. So wirken beispielsweise in den Sozialwissenschaften datenschutzrechtliche Aspekte auf die Publikationsmöglichkeiten und -strategien. Mit Blick auf die kompetitive Forschungslandschaft mag es dem einzelnen Wissenschaftler außerdem notwendig erscheinen, die Veröffentlichung von Daten nur eingeschränkt oder zeitverzögert anzustreben.

Ziel der folgenden übergreifenden Darstellung ist es, einen Überblick auf die gegenwärtige Landschaft der Forschungsdatenpublikation zu schaffen und damit eine Grundlage für zukünftige Diskussionen zu bieten. Hierbei soll vor allem herausgestellt werden, wie die wissenschaftlichen Daten in dem traditionellen Publikationsprozess eingebunden sind.

Mit Blick auf die Naturwissenschaften können drei Publikationsmodelle unterschieden werden[1].

- Die Veröffentlichung von Forschungsdaten als eigenständiges Objekt in einem Datenrepositorium
- Die Veröffentlichung von Forschungsdaten mit einer textuellen Dokumentation
- Die Veröffentlichung von Forschungsdaten als Anreicherung einer interpretativen Text-Publikation.

Alle Modelle setzen bestimmte Rahmenbedingungen in Bezug auf die genutzten Infrastrukturen voraus, insbesondere in Bezug auf eine vertrauenswürdige Publikationsumgebung, die beispielsweise Standards der Langzeitarchivierung berücksichtigen.

Die unterschiedlichen Modelle ziehen unter Umständen unterschiedliche Zeitpunkte der Forschungsdatenpublikation nach sich, insbesondere in Bezug zum interpretativen Artikel. Dieser Aspekt soll hier nicht im Detail ausgeführt werden, er kann jedoch bei der Wahl des Publikationsmodells entscheidend sein.

3.2.3.1 Eigenständige Veröffentlichung im Forschungsdaten-Repositorium

Ein Publikationsmodell ist die Veröffentlichung von wissenschaftlichen Daten in einem Forschungsdaten-Repositorium. Das Repositorium ermöglicht es, For-

1. Darüber hinaus gibt es sicher weitere individuelle Herangehensweisen, auf die in diesem Papier nicht eingegangen werden kann.

schungsdaten zeitlich und örtlich unabhängig von einer interpretativen Veröffentlichung zu veröffentlichen. Die Daten werden als eigenständiges Objekt in ein Datenrepositorium abgelegt. Je nach Disziplin kann dies z. B. nach einer automatisierten Prozessierung aus einem „System“ heraus erfolgen. Dieses Publikationsmodell ist in vielen naturwissenschaftlichen Disziplinen verbreitet. Je nach Disziplin und Organisationstruktur in den Disziplinen variieren die Anforderungen an die Standardisierung und Bereitstellung. Dieses Modell kann auch als Baustein verstanden werden, welches – Standardisierung und Interoperabilitiät vorrausgesetzt – mit anderen Arbeitsabläufen und Produkten der Wissenschaft kombiniert werden kann, siehe dazu auch die Modelle in 3.2.3.2 und 3.2.3.3. Beispielhaft seien hier vier Repositorien für dieses Publikationsmodell genannt, die die Diversität und die Möglichkeiten des Modells widerspiegeln:

- PANGAEA (http://www.pangaea.de): In diesem Datenrepositorium werden geowissenschaftliche Forschungsdaten veröffentlicht. Dank der Nutzung des *Digital Object Identifier* (DOI) wird eine persistente Adressierung der Daten ermöglicht. Komplexere Forschungsdatensätze können als „*Collection*“ publiziert werden.
- GenBank (http://www.ncbi.nlm.nih.gov/genbank): In dieser Datenbank werden Forschungsdaten nach den disziplinären Regeln der Genforschung publiziert. Eine „*accession number*“ dient zur persistenten Adressierung. Nach Vorgaben der *Editorial-Policies* vieler lebenswissenschaftlicher Zeitschriften muss dieser bei einer Interpretation der Daten in einem Aufsatz angegeben werden.
- Integrated Ocean Observing System (http://www.ioos.gov): Hier werden automatisch erfasste Messdaten der marinen Umweltforschung automatisiert aus dem System heraus veröffentlicht.
- Dryad (http://datadryad.org): Dieses Datenrepositorium der Biologie vergibt DOIs für die veröffentlichten Datensätze. Dryad konzentriert sich auf Datensätze, die Grundlage einer interpretativen Publikation sind. Es gibt eine Zitierempfehlung für den Datensatz, welche zusätzlich zu einer Zitierung des interpretativen Artikels erfolgt.

3.2.3.2 Veröffentlichung mit einer textuellen Dokumentation

Im Hinblick auf die Nachnutzung und Nachvollziehbarkeit von Forschungsdaten benötigen diese oftmals eine erweiterte Dokumentation, die zwar keine Interpretation, aber einen Kontext zur Erhebung der Daten liefert und Interessierten eine Nachnutzung ermöglicht. Vor diesem Hintergrund werden neuere Konzepte entwickelt. So werden z. B. Forschungsdaten mit einer begutachteten Dokumentation publiziert.

Im Folgenden sollen zwei Beispiele dieser Publikationsstrategie vorgestellt werden, bei denen Forschungsdaten mit einer nicht-interpretativen Dokumentation angereichert werden (siehe Abbildung 1).

Abb. 1: Die Publikation von Forschungsdaten mit einer erweiterten (begutachteten) Dokumentation und ggf. weitere Materialien, welche die Nachnutzung ermöglichen.Weitere Service können an die Publikation gekoppelt werden. Zusätzliche Analysen und (mehrere) interpretative Artikel können auf den Forschungsdaten aufbauen.

Beispiel: Earth System Science Data (ESSD)[2]

Die *Open-Access*-Zeitschrift *Earth System Science Data* (ESSD) widmet sich der erweiterten und qualitätsgesicherten Dokumentation von geowissenschaftlichen Forschungsdaten. Die Zeitschrift wird von *Copernicus Publications* verlegt und nutzt ein zweistufiges *Peer-Review*-Verfahren unter der Nutzung des „*Public Peer-Reviews*“ (Pöschl, 2010) um einen Datensatz und dessen Dokumentation zu veröffentlichen und zu begutachten.

Der Pubklikationsprozess in ESSD ist wie folgt organisiert: Ein Publizierender verfasst eine Dokumentation eines Datensatzes. In dieser beschreibt er beispielsweise die verwendeten Instrumente und angewandte Verfahren bei der Bearbeitung der Daten. Diese Beschreibung, die wie ein klassischer Aufsatz gestaltet wird, reicht er dann unter Nennung eines Persistent Identifier des Datensatzes bei der Zeitschrift ein. Darüber hinaus veröffentlicht er die beschriebenen Daten auf einem frei zugänglichen Daten-Repositorium. Nach einem erfolgreichen Begutachtungsprozess wird dann die Datendokumentation publiziert. Nach diesem Ansatz sind Datenrepositorium und Datenjournal deutlich getrennt. Das Ambiente der Zeitschrift ESSD erinnert stark an die Journale, welche „interpretativen“ Artikel publizieren.

2. http://earth-system-science-data.net/ [Zugriff am 18.08.2011].

Das Datenjournal ist unabhängig von Repositorien, hat jedoch eigene Kriterien für die Akzeptanz der Repositorienwahl festgelegt (Pfeiffenberger, 2011), welche sich nach den gegenwärtigen Standards für vertrauenswürdige Repositorien (z. B. von NESTOR, 2009) richten.

Für die Wissenschaftler ist der Prozess der Publikation ein bekannter, obwohl das zentrale Publikationsobjekt Forschungsdaten und deren Dokumentation in diesem Rahmen neu ist. Der zusätzliche Aufwand einer Einreichung der Publikationsobjekte zu zwei Publikationsplattformen (Zeitschrift und Datenrepositorium) mag für den Wissenschaftler durch den Erhalt eines zitierfähigen Artikels in einem bekannten Journalformat kompensiert werden.

Beispiel: Overlay Journal Infrastructure Meteorological Sciences (OJIMS)[3]

Unter dem Begriff „*Overlay Journal*“ fasst man Publikationsmodelle zusammen, welche auf Textpublikationen zurückgreifen, die auf verschiedenen Plattformen *Open Access* publiziert sind. Das Konzept beinhaltet eine thematische Aggregation, meist mit einem weiteren Service kombiniert, z. B. einem Begutachtungsprozess- oder Kommunikationsprozess. Das Konzept kann auch auf die Publikation von Forschungsdaten übertragen werden. Das *Overlay Journal Infrastructure Meteorological Science* (OJIMS) wendet dieses für meteorologische Daten an. Die meteorologischen Daten werden auf einem Forschungsdaten-Repositorium gespeichert, die Dokumentation in einem Dokumentenrepositorium. Beides wird anschließend von Fachwissenschaftlern begutachtet. Auf Basis dieser qualitätsgesicherten Publikation können dann Interpretationen der Daten vorgenommen werden.

3.2.3.3 (Interpretative) Publikation plus Forschungsdaten

Traditionell werden wissenschaftliche Ergebnisse in einer Textpublikation beschrieben und diskutiert. Über Disziplingrenzen hinweg lässt sich dabei seit Jahrhunderten eine ähnliche Artikelstruktur verfolgen, die auf die Darstellung eines Ergebnisses fixiert ist. Neben dem Text dienen Tabellen und Abbildungen der Beschreibung des dargestellten Sachverhaltes. Zusätzliche Materialien werden oft in den Anhang dargestellt (siehe Abbildung 2). Mit der Digitalisierung der Wissenschaftskommunikation und der Umstellung auf elektronische Zeitschriften können diese Materialien, z. B. Messreihen im CSV-Format oder Proteinsequenzen im FASTA-Format nachnutzbar und nachprüfbar publiziert werden. Die Publikation solcher Daten, die insbesondere zur Unterstützung einer interpretativen Publikation dienen, wird heute insbesondere auf den folgenden Wegen umgesetzt.

3. http://proj.badc.rl.ac.uk/ojims [Zugriff am 18.08.2011].

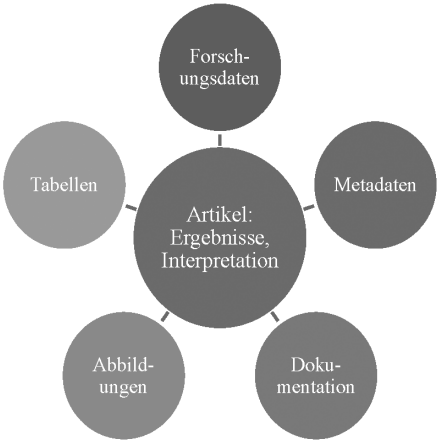

Abb. 2: Die Publikation von Forschungsdaten als Anhang einer interpretativen Publikation. Dieser wird beispielsweise in einem Journal publiziert und die zugrunde liegenden Forschungsdaten werden zusammen mit dem Artikel bereitgestellt. Im Anhang zu dem interpretativen Artikel finden sich traditionell auch oftmals Abbildungen, Tabellen und detailliertere Angaben zur Methodik.

Die Daten werden auf der gleichen Plattform wie auch die Textpublikation veröffentlicht, aber nicht individuell persistent und standardisiert adressiert. Dies kann auf einer Verlags- oder auch Repositorienplattform sein, welche vorwiegend für Textpublikation ausgelegt ist. Die Daten sind somit Bestandteil des Artikels.

In einigen Disziplinen werden Forschungsdaten als eigenständiges Objekt, das referenzierbar ist, auf der Publikationsplattform des textuellen Artikels gespeichert. Dank der Referenzierbarkeit, die über eine persistente Adressierung gegeben ist, kann der Datensatz zitiert werden. In dem interpretativen Artikel wird dann diese Identifizierung angegeben.

Der Datensatz kann in diesem Modell auch in einem externen Datenrepositorium (siehe 3.2.3.1) zeitlich parallel zu dem interpretativen Artikel publiziert werden. Dies bedeutet für den Wissenschaftler, dass dieser den Forschungssdatensatz und den interpretativen Artikel in zwei unterschiedlichen Orten einreicht. Eine koordinierte Zusammenarbeit der involvierten Infrastruktureinrichtungen, oft Verlage und Datenrepositorien, wie sie in 3.2.3.1. beschrieben worden sind, ist dabei wichtig, um zusätzlichen Aufwand auf Seiten der Wissenschaftler zu vermeiden und um den Anforderungen einer vertrauenswürdigen Langzeitarchivierung zu entsprechen. In der Nachnutzung können Textpublikation und Forschungsdaten individuell genutzt und zitiert werden (siehe z. B. die Zitierempfehlung des Datenrepositoriums Dryad). Kooperationen zwischen den

Publikationsagenten Verlag / Journal und Datenrepositorium lassen sich bereits in einigen Disziplinen beobachten. Es sollen hier zwei Beispiele vorgestellt werden.

- PANGAEA und Elsevier: Der Verlag Elsevier und das Datenrepositorium PANGAEA sind eine Kooperation eingegangen (Elsevier, 2010), die erlaubt, dass Leser von Artikeln auf Science Direct direkten Zugang zu Forschungsdaten haben, welche einem geowissenschaftlichen Artikel zugeordnet sind und auf PANGAEA abgelegt wurden. Der Link zu dem Datensatz ist prominent auf der Artikelseite platziert und führt den Leser direkt zum Datensatz.
- Dryad und verschiedene Verlage / Journale: Das Datenrepositorium Dryad kooperiert mit verschiedenen Verlagen und Journalen, insbesondere aus den biologischen Disziplinen. Das Dryad Konsortium hat eine „*Joint Data Archiving Policy*“ entwickelt, welche die Ablage von Forschungsdaten für Einreichungen in den Partnerjournalen regelt. Es wird explizit von Forschungsdaten gesprochen, welche Grundlage der Interpretation im Artikel sind: „*data supporting the results in the paper*“ (Datadryad, 2010).

Einen interessanten Überblick über unterschiedliche, disziplinspezifische Ausprägungen dieses Modells sind in den geförderten SURF-Projekten zu finden, welche an der disziplinspezifischen Umsetzung der sogenannten „*enhanced publications*“ arbeiten (SURF Foundation, 2010). SURF definiert „*enhanced publications*“ als „a publication – usually a text – that has been enhanced with additional material [...] The supplementary material may consist, for example, of research data, illustrative images, metadata sets, or post-publication data such as comments or rankings.“ Sechs disziplinspezifische Projekte werden von SURF seit Januar 2011 gefördert, u.a. aus der Ökonomie und der Linguistik.

3.2.4 Fazit

Die Entwicklung von Publikationsmodellen für Forschungsdaten findet sich aktuell in einer dynamischen Entwicklung, was eine Typisierung der unterschiedlichen Ansätze schwierig macht.

In dieser Übersicht wurden drei Modelle der Publikation von Forschungsdaten beschrieben. Im erstgenannten Modell werden Forschungsdaten zitierfähig auf einem Repostorium veröffentlicht. In einem zweiten Modell werden Forschungsdaten mit einer „erweiterten“ Dokumentation einem Begutachtungsprozess unterzogen. Zusätzliche und qualitätsgesicherte Kontextinformationen erleichtern dann die Nachnutzung. Ein bereits in zahlreichen Disziplinen umgesetztes Modell ist die Forschungsdatenpublikation als Anhang zu einem interpretativen Artikel. Hierbei kann der Datensatz z. B. in einem externen Datenrepositorium abgelegt werden und dann in dem Artikel referenziert werden.

Die weitere Entwicklung der Publikationsmodelle wird stark von den Publikationskulturen der einzelnen Disziplinen abhängig sein. Aufgrund der hohen disziplinspezifischen Charakteristika der Forschungsdaten sind angepasste Lösungen, die den Anforderungen der Disziplinen entsprechen, nötig. Der Kooperation zwischen Wissenschaft, Infrastruktureinrichtungen und Verlagen kommt dabei eine entscheidende Rolle zu.

Literaturhinweise

AGU Publications Committee, 1993. *Policy on Referencing Data in and Archiving Data for AGU Publications*. (Revised: March 1994, Dec. 1995, Oct. 1996) Online: http://www.agu.org/pubs/authors/policies/data_policy.shtml [Zugriff am 22.02.2011].

Bermuda Principles, o. J. *Policies on Release of Human Genomic Sequence Data*. (Last modified: 29.10.2003) Online: http://www.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml [Zugriff am 22.02.2011].

Borgman, C. L., 2010. Research Data: Who will share what, with whom, when, and why? *5th China-North American Library Conference*. Peking, China 8.-12. Sept. 2010. Online: http://works.bepress.com/borgman/238/ [Zugriff am 22.02.2011].

Dallmeier-Tiessen, S. & Pfeiffenberger, H., 2008. Peer Reviewed Data Publication in Earth System Sciences. In: C. Puschmann & D. Stein, D., Hrsg. 2008. *Towards Open Access Scholarship*. Düsseldorf: Düsseldorf University Press, S.77–84.

DataDryad, 2010. *Joint Data Archiving Policy (JDAP)*. Online: http://datadryad.org/jdap [Zugriff am 22.02.2011].

DFG (Deutsche Forschungsgemeinschaft), 1998. *Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“*. Denkschrift. Weinheim: Wiley-VCH.

DFG (Deutsche Forschungsgemeinschaft), 2010. *Merkblatt für Anträge auf Sachbeihilfen mit Leitfaden für die Antragstellung.*(DFG-Vordruck 1.02 – 8/10) Online: http://www.dfg.de/download/programme/sachbeihilfe/antragstellung/1_02/1_02.pdf [Zugriff am 18.08.2011].

Elsevier, 2010. *„Elsevier and PANGAEA Take Next Step in Connecting Research Articles to Data“*. Online: http://www.elsevier.com/wps/find/authored_newsitem.cws_home/companynews05_01616 [Zugriff am 22.02.2011].

Hey, T. Tansley, S. & Tolle, K., 2009. Jim Gray on eScience: a transformed scientific method. In: S. Tansley & K. Tolle, Hrsg. 2009. *The Fourth Paradigm: Data-Intensive Scientific Discovery.* Redmond, Wash.: Microsoft Research, S. XVII–XXXI. Online: http://www.fourthparadigm.org [Zugriff am 18.08.2011].

Marshall, E., 2001. Bermuda Rules: Community Spirit, With Teeth. *Science*,291(5507), S. 1192. DOI*:* 10.1126/science.291.5507.1192

NSF (National Science Foundation), 2011. *National Grant Proposal (GPG)*. Chapter II – Proposal Preparation Instructions, C. 2.j. Data Management. Online: http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp#dmp, 2011 [Zugriff am 22.02.2011].

NESTOR – Kompetenznetzwerk Langzeitarchivierung, 2009. *NESTOR-Kriterien – Kriterienkatalog vertrauenswürdige digitale Langzeitarchive.* (Version II). Frankfurt am Main: NESTOR.

OECD (Organisation for Economic Co-operation and Development), 2007. *Principles and Guidelines for Access to Research Data from Public Funding*. Online: http://www.oecd.org/dataoecd/9/61/38500813.pdf [Zugriff am 20.02.2011].

Pfeiffenberger, H. & Carlson, D., 2011. Earth System Science Data (ESSD) – A Peer Reviewed Journal for Publication of Data. *D-Lib Magazine*, (17)1/2.

PLoS ONE, o. J. *PLoS ONE Editorial and Publishing Policies.* Sharing of Materials, Methods, and Data. Online: http://www.plosone.org/static/policies.action#sharing [Zugriff am 22.02.2011].

Pöschl, U.: Interactive Open Access Publishing and Peer Review: The Effectiveness and Perspectives of Transparency and Self-Regulation in Scientific Communication and Evaluation. In: Liber Quarterly 19, 2010, Nr. 3/4.

SURF Foundation, 2010. *Enhanced Publication.* Online: http://www.surffoundation.nl/nl/publicaties/Documents/SHAREflyer_verrijkte_publicatie_pdfversie_def_ENG.pdf [Zugriff am 22.02.2011].

Zacharias, M. C., 2010. *Scientists Seeking NSF Funding Will Soon Be Required to Submit Data Management Plans.* Online: http://www.nsf.gov/news/news_summ.jsp?cntn_id=116928 [Zugriff am 18.08.2011].