



Heike Neuroth, Stefan Strathmann, Achim Oßwald,
Regine Scheffel, Jens Klump, Jens Ludwig [Hrsg.]

Langzeitarchivierung von Forschungsdaten

Eine Bestandsaufnahme

Kapitel 1

Langzeitarchivierung von Forschungsdaten

Einführende Überlegungen

Langzeitarchivierung von Forschungsdaten

Eine Bestandsaufnahme

Heike Neuroth, Stefan Strathmann, Achim Oßwald,
Regine Scheffel, Jens Klump, Jens Ludwig [Hrsg.]

Langzeitarchivierung von Forschungsdaten

Eine Bestandsaufnahme



Förderkennzeichen: 01 DL 001 B

Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme
hg. v. Heike Neuroth, Stefan Strathmann, Achim Oßwald, Regine Scheffel, Jens Klump,
Jens Ludwig
im Rahmen des Kooperationsverbundes nestor – Kompetenznetzwerk Langzeitarchivierung
und Langzeitverfügbarkeit digitaler Ressourcen für Deutschland
nestor – Network of Expertise in Long-Term Storage of Digital Resources
<http://www.langzeitarchivierung.de/>

Kontakt: editors@langzeitarchivierung.de
c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen,
Dr. Heike Neuroth, Forschung und Entwicklung, Papendiek 14, 37073 Göttingen

Die Herausgeber danken Anke Herr (Lektorat) und Sonja Neweling (Redaktion) sowie Martina Kerzel und Lajos Herpay (Gestaltung und Montage) für ihre unverzichtbare Unterstützung bei der Fertigstellung des Handbuchs.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter
<http://www.d-nb.de/> abrufbar.

Die Inhalte des Buches stehen auch als Onlineversion
(<http://nestor.sub.uni-goettingen.de/bestandsaufnahme/>)
sowie über den Göttinger Universitätskatalog
(<http://www.sub.uni-goettingen.de>) zur Verfügung.
Die URN lautet: <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>.

Die digitale Version dieses Werkes steht unter einer Creative Commons Namensnennung-
Nicht-kommerziell-Weitergabe unter gleichen Bedingungen 3.0 Unported Lizenz.



Einfache Nutzungsrechte liegen beim Verlag Werner Hülsbusch, Boizenburg.
© Verlag Werner Hülsbusch, Boizenburg, 2012
<http://www.vwh-verlag.de>
In Kooperation mit dem Universitätsverlag Göttingen.

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen,
Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und
als solche den gesetzlichen Bestimmungen unterliegen.

Druck und Bindung: Kunsthaus Schwanheide
Printed in Germany – Als Typoskript gedruckt –

ISBN: 978-3-86488-008-7

1 Langzeitarchivierung von Forschungsdaten

Einführende Überlegungen

Achim Oßwald, Regine Scheffel, Heike Neuroth

Spätestens seit die Medien darüber berichtet haben, dass die NASA Daten des ersten bemannten Mondflugs nur mit hohem Aufwand wieder herstellen konnte,¹ wurde erkennbar, dass erhebliche Anstrengungen notwendig sind, um digitale Daten aus der Wissenschaft für die Zukunft zu sichern. Weitere Datenpannen großen Ausmaßes² bestätigen diese Notwendigkeit auch für andere Bereiche des gesellschaftlichen Lebens. Darüber hinaus gibt es immer wieder Belege für die bewusste Manipulation von Forschungsdaten durch Wissenschaftler.³

Für den Wissenschaftsbetrieb ist der verlässliche und langfristige Zugriff auf Forschungsdaten aus mehreren Gründen notwendig: Der Skandal z.B. um den Zellbiologen Tae Kook Kim⁴ hat deutlich gemacht, wie wichtig es ist, Forschungsdaten, auf denen aktuelle wissenschaftliche Veröffentlichungen beruhen, verfügbar und damit auch verifizierbar zu halten. Digitale Forschungsdaten – heute unverzichtbare Grundlage wissenschaftlicher Erkenntnis – sind aber vielfach nicht mehr reproduzierbar und damit im Falle eines Verlustes unwiderruflich verloren, d.h. auch nicht mehr verifizierbar. Messdaten aus dem Bereich der Klimaforschung aus den letzten Jahrzehnten zeigen dies z.B. eindrücklich. In diesen Fällen macht erst die Archivierung und Langzeitverfügbarkeit die Verifizierbarkeit, Interpretierbarkeit und Nachnutzung der einmal erhobenen Forschungsdaten möglich. Dabei ist zu beobachten, dass sich die Formen der Nachnutzung mit den erweiterten Möglichkeiten des Zugriffs ausdifferenzieren: zum einen eröffnet die Integration von digitalen Daten in neue fachwissenschaftliche Kontexte der Wissenschaft neue Chancen,

1 Schmundt (2000); Hammerschmidt (2002).

2 Vgl. z.B. o.V. (2007).

3 Vgl. z.B. o.V. (2010).

4 Vgl. Kennedy; Alberts (2008).

indem alte Forschungsfragen neu beantwortet oder gänzlich neue Forschungsfragen generiert werden. Zum anderen werden Langzeitstudien wie zum Beispiel in der Klimaforschung oder in den Sozialwissenschaften erst möglich. In der Astronomie zum Beispiel⁵ wurde seit Ende des 19. Jahrhunderts die (analoge) Fotografie genutzt, um Stern Daten dauerhaft zu speichern. Zu den umfangreichsten Datensammlungen gehören das Archiv des Harvard College Observatory⁶ mit über 500.000 Fotoplatten, die in einem Zeitraum von mehr als 100 Jahren bis 1989 aufgenommen wurden. Ein anderes Beispiel ist das Archiv der Sternwarte Sonneberg⁷, das etwa 300.000 Fotoplatten aus 70 Jahren umfasst, mit denen mehr als 10.000 veränderliche Sterne entdeckt wurden. Diese enormen Datenarchive werden nach und nach digitalisiert, um sie der Nachwelt zu erhalten und mit Computertechniken auswerten zu können. Insbesondere für Helligkeitsveränderungen in Zeiträumen von vielen Dutzenden von Jahren und für Positionsänderungen von Sternen sind sie eine unverzichtbare Quelle.

Auch die interdisziplinäre Nutzung von Daten wird erst durch den freien Zugang und die Referenzierbarkeit von Forschungsdaten ermöglicht. Eine neue Form der Nachnutzung entwickelt sich in den USA mit dem Trend zum *Crowdsourcing*⁸, bei dem die Öffentlichkeit oder eine definierte Teilmenge der (Fach-) Öffentlichkeit (z.B. Diplomanden) an der Generierung oder qualitativen Anreicherung von Forschungsdaten beteiligt werden. Ein Beispiel für *Citizen Science*⁹ oder *Crowdsourcing*, bei dem interessierte Laien mit in den Forschungsprozess eingebunden werden, ist das Projekt *Galaxy Zoo*¹⁰: Moderne Himmelskartierungen erstellen unzählige Aufnahmen von Galaxien bzw. weit entfernten Milchstraßen. Die Galaxienformen sind so vielfältig und komplex, dass es bis heute keine gute computerisierte Klassifizierung gibt. Im Juli 2007 gingen amerikanische Astrophysiker mit diesem Thema deshalb an die Öffentlichkeit.

5 Freundliche Mitteilung von Prof. Wambsganß, Astronomisches Rechen-Institut (ARI) am Zentrum für Astronomie der Universität Heidelberg (ZAH).

6 Vgl. Harvard College Observatory (2011).

7 Vgl. Sternwarte Sonneberg (2011).

8 Vgl. o.V. (2011b).

9 Vgl. o.V. (2011a).

10 Vgl. Galaxy Zoo (2011).

Sie luden interessierte Sternenfreunde ein, bei der Klassifizierung dieser Galaxien mitzuwirken. Dazu wurden Trainingssequenzen angeboten, so dass die neuen Teilnehmer die Kriterien zur Klassifizierung lernen konnten. Als strukturell ähnlich gelagertes Beispiel sei aus den Geisteswissenschaften das Projekt *Collaborative Manuscript Transcription*¹¹ genannt.

Im Sinne der Langzeitarchivierung (LZA) geht es also darum, einerseits Forschungsdaten – zum Teil auch als eigenständige Publikation¹² – langfristig digital zur Verfügung zu stellen und damit verifizierbar, interpretierbar und nachnutzbar zu machen und andererseits Forschungsdaten auf der Basis von Forschungsinfrastrukturen miteinander zu vernetzen und so insbesondere die potentielle Nachnutzung auch interdisziplinär zu erhöhen. Damit rückt zugleich eine neue Vision einer Forschungsumgebung in den Fokus, die die *High Level Expert Group on Scientific Data*, eine Expertengruppe der Europäischen Kommission, im Oktober 2010 als Vision 2030 für Forschungsdaten formuliert hat:

“Our vision is a scientific e-infrastructure that supports seamless access, use, re-use, and trust of data. In a sense, the physical and technical infrastructure becomes invisible and the data themselves become the infrastructure – a valuable asset, on which science, technology, the economy and society can advance.”¹³

Die Realisierung dieser Vision ist jedoch noch mit einer Reihe offener Fragen und ungelöster Probleme verbunden. Das beginnt schon bei dem Begriff *Forschungsdaten* selbst: Was sind Forschungsdaten? Dies können zum Beispiel Messdaten aus Instrumenten wie einem Teleskop sein, Rohdaten aus einem Massenspektrometer, aber auch digitale Karten oder Volltexte z.B. zur Erstellung kritischer Editionen, d.h. der Begriff Forschungsdaten ist immer in Bezug zur jeweiligen Fachdisziplin zu setzen. Ebenso sind Anforderungen an das Management und die Langzeitverfügbarkeit (LZV) von Forschungsdaten immer auf generische sowie fachspezifische Aspekte und Lösungen hin ausdifferenzieren.

Bislang gibt es (nicht nur) in Deutschland noch keine Definition von digitaler Langzeitarchivierung, die sich eines breiten Konsenses erfreut. Auch das deutsche Kompetenznetzwerk zur digitalen Langzeitarchivierung

11 Vgl. *Collaborative Manuscript Transcription* (2011).

12 Vgl. z.B. PANGAEA (2011a).

13 High Level Expert Group on Scientific Data (2010), S.4.

nestor¹⁴, das sich seit Jahren intensiv mit diesem Themenbereich beschäftigt, bietet auf seiner Homepage keine Definition an. Daher soll an dieser Stelle auf die Einführung des nestor-Handbuchs¹⁵ verwiesen werden:

„Langzeitarchivierung meint in diesem Zusammenhang mehr als die Erfüllung gesetzlicher Vorgaben über Zeitspannen, während der steuerlich relevante tabellarisch strukturierte Daten verfügbar gehalten werden müssen. ‚Langzeit‘ ist die Umschreibung eines nicht näher fixierten Zeitraumes, währenddessen wesentliche, nicht vorhersehbare technologische und soziokulturelle Veränderungen eintreten; Veränderungen, die sowohl die Gestalt als auch die Nutzungssituation digitaler Ressourcen in rasanten Entwicklungszyklen vollständig umwälzen können. Es gilt also, jeweils geeignete Strategien für bestimmte digitale Sammlungen zu entwickeln, die je nach Bedarf und zukünftigem Nutzungsszenarium die langfristige Verfügbarkeit und Nachnutzung der digitalen Objekte sicherstellen. ‚Langzeit‘ bedeutet für die Bestandserhaltung digitaler Ressourcen nicht die Abgabe einer Garantierklärung über fünf oder fünfzig Jahre, sondern die verantwortliche Entwicklung von Strategien, die den beständigen, vom Informationsmarkt verursachten Wandel bewältigen können.“^{16,17}

Unter Langzeitarchivierung verstehen wir hier die für den jeweiligen Kontext definierte Zeitspanne der Bewahrung digitaler Objekte auch über grundlegende technologische und soziokulturelle Wandlungsprozesse hinaus. Mit ihr wird ermöglicht, den Zugriff auf und die Nachnutzung von Forschungsdaten auch in der Zukunft zu sichern.

Die Herausforderungen, die sich daraus ableiten, liegen auf der Hand: Wir können nicht alle Forschungsdaten aufbewahren. Was aber sind dann die Auswahlkriterien und wer legt sie fest? Wer kann heutzutage sicher abschätzen, was auch für zukünftige (Forscher)Generationen von Interesse ist? Wie gehen wir mit Forschungsdaten um, die sich nicht mehr reproduzieren lassen (z.B. die angesprochenen Klimadaten oder Himmelsbeobachtungen)? Deutlich ist, dass die reine *Bitstream-Preservation*¹⁸, also

14 Vgl. nestor (2011a).

15 Vgl. die gedruckte Ausgabe 2.0 des nestor Handbuchs von 2009 sowie die aktuelle Online-Version 2.3 von 2010.

16 Liegmann; Neuroth (2010), Kap. 1:5f.

17 In diesem Zusammenhang stellt sich die Frage, ob die ursprünglich von Schwens und Liegmann 2004 publizierte Erläuterung von Langzeitarchivierung und Langzeitverfügbarkeit auch von den wissenschaftlichen Communities angenommen werden kann. Vgl. Schwens; Liegmann (2004), S. 567.

18 Vgl. Ullrich (2010), Kap. 8:3-8:9.

die alleinige Bewahrung der Bits und Bytes des *physical object*¹⁹, hierfür allenfalls ein erster Schritt sein kann. Langzeitverfügbarkeit, also die zukünftige Interpretierbarkeit und Nutzbarkeit von wissenschaftlichen Daten, stellt ungleich höhere Anforderungen, weil künftige technische Nutzungsumgebungen nicht prognostizierbar sind. Deshalb müssen digitale Objekte, die in ein Langzeitarchiv²⁰ eingeliefert werden, durch Metadaten beschrieben werden. Bewahrt und (standardisiert) dokumentiert werden muss auch, unter welchen technischen und organisatorischen Rahmenbedingungen die Daten entstanden sind. Nur dann besteht eine Chance, dass ihre Nutzung in Zukunft – ggf. auf der Grundlage von Emulation²¹ oder Migration²² – möglich sein wird.²³ Doch bereits auf kürzere Sicht werden schon beschreibende (deskriptive), technische und administrative Metadaten benötigt. Dies zeigen u.a. die Beispiele aus dem *Factsheet Keeping Research Data Safe* (KRDS)²⁴, einer Zusammenfassung von zwei Studien zu den Kosten der Langzeitarchivierung von Forschungsdaten.²⁵ Selbst Forschungsergebnisse von Studien, die erst vor einigen Jahre abgeschlossen wurden, konnten von den beteiligten Wissenschaftlern nicht weitergenutzt werden, weil z.B. die Erhebungsmethoden nicht detailliert genug dokumentiert wurden.²⁶ Dies gilt erst recht, wenn Forschungsdaten über grundlegende soziokulturelle Wandlungsprozesse hinaus langfristig nutzbar erhalten werden sollen: So werden z.B. alte Kirchenbücher heute unter Genderaspekten analysiert, ein Aspekt, der damals sicherlich nicht angedacht wurde. Um heutige Verwaltungsunterlagen und Datenbanken, die vergleichbare Daten beinhalten, für künftige Forschungsfra-

19 Thibodeau unterscheidet bei einem digitalen Objekt die Ebene des *conceptual object*, das als bewahrungswürdig eingestuft wurde, das *logical object* der Realisierung in Form von Dateien, die an eine bestimmte Hard- und Softwareumgebung gebunden sind, sowie das *physical object* des reinen Bitstreams (vgl. Thibodeau (2002)).

20 Hier wird von einem Langzeitarchiv nach dem OAIS-Modell ausgegangen. Vgl. das Ergebnis der Diskussion zur aktuellen Version des Standards unter Reference Model for an Open Archival Information System (vgl. OAIS (2009)) und die sich wiederum hierauf beziehenden Diskussionen. Einen Überblick zu OAIS gibt das Referenzmodell OAIS (vgl. OAIS (2010)).

21 Vgl. Funk (2010a), Kap. 8:16-8:23.

22 Vgl. Funk (2010b), Kap. 8:10-8:15.

23 Unter welchen rechtlichen Bedingungen dies dann zulässig sein wird, ist ebenfalls zu klären.

24 Vgl. Charles Beagrie Ltd; JISC (2010).

25 Vgl. Beagrie; Chruszcz; Lavoie (2008); Beagrie; Lavoie; Woollard (2010).

26 Vgl. Beagrie; Lavoie; Woollard (2010), S. 2.

gen nutzbar zu machen, müssen entsprechende Metadaten mit erfasst, archiviert und verfügbar gemacht werden. Hierbei sollte von unterschiedlichen zukünftigen Nutzungsszenarien und potentiellen Nutzergruppen (sog. *designated communities*) und deren Erwartungen an die Beschreibung der überlieferten Daten in die konzeptionellen Überlegungen der Langzeitarchivierung ausgegangen werden.

Folglich kommt den deskriptiven Metadaten eine besondere Bedeutung zu, insbesondere den systematisch ausdifferenzierten Detailangaben, die Aufschluss geben über die Kriterien für die Auswahl des Untersuchungsgegenstands, die Untersuchungs-, Mess- oder Erhebungsmethoden, ihre Anwendung sowie über die Untersuchungsergebnisse selbst. Die hier vorgelegte Bestandsaufnahme (vgl. Kap. 4 und Folgende) versucht u.a. zu ermitteln, ob es bereits disziplinabhängige oder -übergreifende Konventionen gibt, die im Hinblick auf die Langzeitverfügbarkeit von Forschungsdaten und den Aufbau von deutschlandweiten Forschungsinfrastrukturen relevant sind.

Insgesamt wird deutlich, dass derartiges digitales Kuratieren (*digital curation*)²⁷ von Forschungsdaten im Hinblick auf Langzeitarchivierung und Langzeitverfügbarkeit auch bereits für den aktuellen Forschungsbetrieb Vorteile hat. Die Zugänglichkeit publizierter Forschungsdaten dient der Qualitätssicherung des Wissenschaftsbetriebs und der Erleichterung des wissenschaftlichen Publizierens.²⁸ Als Nebeneffekt erhöht sie das Forschungsniveau und die Produktivität des Forschungsbetriebs. Dies zeigt sich z.B. in ganz pragmatischen Aspekten wie der Sicherstellung der Kontinuität von Forschungsarbeiten über mehrere Forschergenerationen hinweg. Ein weiterer Vorteil der systematischen Dokumentation und Pflege von Forschungsdaten bereits während ihrer Produktion liegt in der langfristigen Kostenersparnis. Die nachträgliche Korrektur fehlerhafter Metadaten kann u.U. um den Faktor 30 teurer sein als die eigentliche Erstellung der Daten selbst.²⁹

27 Der Begriff digitales Kuratieren, eine Übersetzung des englischen *digital curation*, beginnt sich auch im deutschsprachigen Bereich immer stärker durchzusetzen für die systematische Planung, Erstellung, Pflege, Bewertung und ggf. Kassation oder Transformation und Weiterentwicklung digitaler Forschungsdaten und – im weiteren Sinne – aller digitalen Objekte (vgl. DCC (2011b)).

28 Charles Beagrie Ltd; JISC (2010), S. 2.

29 Vgl. ebenda.

Längst haben die Wissenschaftsorganisationen in Deutschland hierauf mit Vorgaben für eine vorsorgliche Datensicherung reagiert. So ist z.B. eine Auflage der Deutschen Forschungsgemeinschaft (DFG) an Projektnehmer, die Daten, auf denen ihre Forschungsergebnisse beruhen, mindestens zehn Jahre verfügbar zu halten.³⁰ Die Allianz der deutschen Wissenschaftsorganisationen³¹ arbeitet daran, die Bereitstellung und Nachnutzung von Forschungsdaten durch die Entwicklung von Standards, Archivstrukturen und Anreizsysteme zu verbessern.³² Auch der Wissenschaftsrat (WR) hat sich in seinen „Übergreifende(n) Empfehlungen zu Informationsinfrastrukturen“³³ vom Januar 2011 diesbezüglich klar positioniert und eine dauerhafte Förderung entsprechender Forschungsinfrastrukturen und Langzeitarchivierungskonzepte eingefordert. Die Identifizierung von Forschungsdaten durch persistente Identifikatoren (z.B. URN³⁴, DOI³⁵, EPIC³⁶) ist dabei ein wesentlicher Schritt u.a. auch zur dauerhaften Zitierfähigkeit dieser Daten und Datensammlungen. Doch die 10-Jahres-Perspektive der DFG ist nur ein Beitrag zur Datensicherung, die Nachnutzung der Forschungsdaten setzt ihre Langzeitarchivierung und Langzeitverfügbarkeit voraus.

Für das Gelingen der Langzeitarchivierung von Forschungsdaten spielen Kooperationen eine zentrale Rolle. Sie finden auf verschiedenen Ebenen statt: Auf der lokalen bzw. institutionellen Ebene, die den Vorteil bietet, dass hier Auswirkungen für die Wissenschaftler am besten einzuschätzen sind und Einflussnahme direkt stattfinden kann. Auf der regionalen, sicher aber auf der nationalen Ebene werden dann institutionell und/oder gesetzlich abgesicherte Maßnahmen angesiedelt sein. Auf der europäischen und internationalen Ebene sind ebenfalls (im Idealfall standardisierte) Strukturen und Verfahren zu etablieren, um der zunehmend global stattfindenden Forschung Rechnung tragen zu können. Disziplinspezifische Daten-

30 Vgl. DFG (1998), S. 12.

31 Vgl. Allianz-Initiative (2011b).

32 Vgl. Allianz-Initiative (2010a).

33 Vgl. Wissenschaftsrat (2011b).

34 Vgl. Schöning-Walter (2010).

35 Vgl. Brase (2010).

36 Vgl. EPIC (2011).

zentren³⁷, die bereits heute effizientes Datenmanagement gewährleisten, können dabei Knotenpunkte eines Langzeitarchivierungsnetzwerks werden. Zusammen bilden sie eine auf die Langzeitverfügbarkeit wissenschaftlicher Forschungsdaten angelegte Langzeitarchivierungsinfrastruktur.

Auch wenn in den vergangenen Jahren bereits umfangreiche Vorarbeiten geleistet und Konzepte für den nachhaltigen Umgang mit Forschungsdaten erarbeitet wurden, so steckt deren Umsetzung immer noch in den Anfängen. Ein wichtiger Faktor scheint dabei zu sein, dass die bisher erprobten Lösungen sich noch nicht gut genug in die wissenschaftliche Arbeit integrieren lassen. Eine Studie der SURFfoundation³⁸ untersuchte die Ergebnisse aus 15 Projekten zum Umgang mit Forschungsdaten, insbesondere welche Anforderungen seitens der Wissenschaftler an Forschungsdateninfrastrukturen gestellt werden und zu erfüllen sind, damit diese die Infrastrukturen für Forschungsdaten nutzen. In der Zusammenfassung der untersuchten Studien werden zwei Rollen unterschieden: der Wissenschaftler als Produzent von Daten und der Wissenschaftler als Konsument von Daten. Dabei zeigt sich, dass die Bedürfnisse dieser beiden Perspektiven fast diametral auseinander liegen: Während der Konsument einen zentralen Zugang mit vielfältigen Kombinationsmöglichkeiten von Daten und Werkzeugen erwartet, wünscht der Produzent eine lokal betreute, maßgeschneiderte Arbeitsumgebung. Hinzu kommt, dass formelle Regelungen, Datenmanagementpläne und deren Überprüfung als Hindernisse wahrgenommen werden. Diese Widersprüche zwischen den Akteuren zu überbrücken, bleibt eine große Herausforderung. Ein wichtiges Anliegen muss daher sein, die Ursachen für diese Ambivalenz genauer zu untersuchen und herauszufinden, wie sie abgebaut werden kann. Eine Möglichkeit könnte zum Beispiel darin bestehen, eine intuitiv nutzbare Infrastruktur zur Verfügung zu stellen oder Anreiz- bzw. Sanktionssysteme zu etablieren und damit die Entwicklung einer neuen Publikationskultur von Forschungsdaten zu beför-

37 Insbesondere die Helmholtz-Gemeinschaft (HGF) betreibt bereits einige fachspezifische Datenzentren, z.B. Deutsches Fernerkundungszentrum am Deutschen Zentrum für Luft- und Raumfahrt, World Data Center for Remote Sensing of the Atmosphere (WDC-RSAT) u.a.

38 Vgl. Eeijen (2011).

dern. Den hier genannten Herausforderungen können sich nur Politik, Wissenschaft und Infrastruktureinrichtungen gemeinsam stellen. Dabei gilt es fachspezifische Besonderheiten und Anforderungen zu berücksichtigen und zu beachten, dass dieser Prozess nur aus den jeweiligen Fachdisziplinen heraus gestartet werden kann. Ein Top-down-Ansatz oder eine Standardlösung für alle Disziplinen werden nicht akzeptiert werden und deshalb wenig erfolgversprechend sein.

In den vergangenen Jahren war die öffentliche Diskussion um die Langzeitarchivierung digitaler Daten – z.B. im Zusammenhang mit Nestor – stark auf einen eher traditionell verstandenen Bereich des kulturellen Erbes fokussiert. Es ist an der Zeit, dass Politik und Öffentlichkeit auch Forschungsdaten als nationales, wissenschaftliches Kulturgut erkennen und zum (Förder-) Gegenstand von bewahrenden Infrastrukturmaßnahmen machen.