Technische Universität Berlin

Organisatorisch-technisches Konzept für eine Forschungsdaten-Infrastruktur in der TU Berlin

– Langfassung mit Erläuterungen, Begriffserklärungen und Links –

AG Forschungsdaten (M. Kuberek)

15. März 2012

Servicezentrum Forschungsdaten und -publikationen · SZ F.

Kontakt:

Monika Kuberek

Universitätsbibliothek der Technischen Universität Berlin

E-Mail: monika.kuberek@tu-berlin.de

Das Konzept steht in einer Kurzfassung und in einer Langfassung online auf der Website des „Servicezentrums Forschungsdaten und –publikationen“ zur Verfügung.

Kurzfassung:
http://www.szf.tu-berlin.de/fileadmin/f33_szf/TUB_Forschungsdaten_Konzept_kurz_20120307.pdf

Langfassung:
http://www.szf.tu-berlin.de/fileadmin/f33_szf/TUB_Forschungsdaten_Konzept_lang_20120315.pdf

DOI: http://dx.doi.org/10.14279/depositonce-6604

Beide Versionen des Konzeptes stehen unter folgender Creative-Commons-Lizenz:
„Attribution-NonCommercial-ShareAlike 3.0 Unported“
http://creativecommons.org/licenses/by-nc-sa/3.0/

Inhalt

Management Summary **4**

1 Ausgangslage und Zielsetzung **4**

1.1 Forschungsdaten in der TU Berlin 5

1.2 Services für die TU-Wissenschaftler/innen und Mehrwert für die TU Berlin 6

2 Inhalt und Rahmen des Konzeptes **6**

2.1 Definition Forschungsdaten 6

2.2 Lebenszyklus von Forschungsdaten (Research Data Lifecycle) 7

2.2.1 Erstellen – Aussortieren/Löschen 7

2.2.2 Speichern – Archivieren 7

2.2.3 Metadaten – Persistent Identifier 8

2.3 Open Access 8

2.4 Langzeitarchivierung 8

2.5 Repositorien 9

3 Organisatorisches Modell: „Servicezentrum Forschungsdaten und -publikationen“ **10**

3.1 Nutzung vorhandener Kompetenzen 11

3.2 Verteilte Zuständigkeiten und strukturierte Workflows 12

3.3 Organisationsstruktur 14

4 Technische Infrastruktur **14**

5 Umsetzung des Konzepts: Projekt „Forschungsdaten“ **15**

6 Perspektive: Gesamtnachweis der in der TU Berlin erbrachten Forschungsleistungen **16**

6.1 Nachweis und Recherche der Forschungsergebnisse 16

6.2 Perspektive Gesamtnachweis 16

Anlage: Glossar AG Forschungsdaten (D. Otto) **19**

A-Z 20

Positionen zum Umgang mit Forschungsdaten und Forschungsergebnissen 24

Management Summary

„Unter Forschungsdaten sind … digitale und elektronisch speicherbare Daten zu verstehen, die im Zuge eines wissenschaftlichen Vorhabens z.B. durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.“ (DFG 2010) [1]

Das vorliegende organisatorisch-technische Konzept hat zum Ziel, an der TU Berlin eine Forschungsdaten-Infrastruktur aufzubauen, die die langfristige Sicherung und Verfügbarkeit der hier entstandenen Forschungsdaten und der darauf basierenden Forschungspublikationen[2] gewährleistet.

Um die Forschungsdaten-Infrastruktur organisatorisch in der TU Berlin zu verankern und die Nachhaltigkeit der getroffenen Maßnahmen sicherzustellen, sieht das Konzept vor, dass Universitätsbibliothek (UB), IT-Service-Center (tubIT) und Forschungsabteilung in einem „Servicezentrum Forschungsdaten und -publikationen“ kooperieren. Zur Unterstützung der Wissenschaftler/innen stellt das Servicezentrum ein umfassendes Dienstleistungs- und Beratungsangebot zu Forschungsdaten und -publikationen bereit. Es wird eine technische Infrastruktur aufgebaut, die dem großen Spektrum unterschiedlicher Fachdisziplinen an der TU Berlin gerecht wird und in der Lage ist, Forschungsdaten vieler Disziplinen zu verwalten und zu speichern. Dabei werden Forschungsdaten und -publikationen als Einheit angesehen, die gegenseitig aufeinander referenzieren; beide werden im institutionellen Forschungsdaten-Repositorium der TU Berlin gespeichert. In dem Konzept wird berücksichtigt, dass eine Reihe von Fach-Communities in den vergangenen Jahren weltweit disziplinspezifische Forschungsdaten-Infrastrukturen aufgebaut hat, an denen sich auch Wissenschaftler/innen der TU Berlin beteiligen. Zur Dokumentation der von Wissenschaftler/innen der TU Berlin erbrachten Forschungsleistungen ist perspektivisch ein Gesamtnachweis – inkl. Link zu den Forschungsdatensets und Volltexten – über die verteilten Forschungsdaten-Repositorien (institutionell und weltweit) hinweg möglich.

Das Konzept soll im Rahmen eines Projektes ab der ersten Jahreshälfte 2012 umgesetzt werden.

Erstellt wurde das Konzept zum Umgang mit Forschungsdaten in der TU Berlin von einer Arbeitsgruppe, die – unter Federführung der UB und mit Beteiligung von tubIT und Forschungsabteilung – im September 2011 von der Kanzlerin eingesetzt wurde.[3]

1 Ausgangslage und Zielsetzung

Im letzten Jahrzehnt ist weltweit die Erkenntnis gewachsen, dass Forschungsdaten eine unschätzbare Ressource und einen Wert an sich darstellen, den es zu erhalten, disziplinübergreifend bereitzustellen und nachzunutzen gilt. Diese Erkenntnis ging einher mit neuen Technologien zur Datenspeicherung und -archivierung, der stetigen Ausbreitung des Internet in alle Lebensbereiche, der wachsenden Vernetzung und kollaborativen Arbeitsweise der wissenschaftlichen Community und dem unaufhaltsamen Anwachsen der Forschungsdatenmenge. Neben dem Aspekt der Verifizierung von Forschungsergebnissen ist insbesondere der Aspekt der Nachnutzung in den Fokus gerückt. Die Allianz der deutschen Wissenschaftsorganisationen sieht angesichts des hohen Aufwandes für die Gewinnung von Forschungsdaten *„einen dringenden Handlungsbedarf hinsichtlich der systematischen Sicherung, Archivierung und Bereitstellung … für die Nachnutzung durch Dritte“*.[4]

[1] www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_forschungsdaten_1001.pdf , S. 1. Weitere Begriffe für Forschungsdaten: Rohdaten, Primärdaten, Forschungsprimärdaten. (s.a. Anlage: Glossar AG Forschungsdaten (D. Otto); Stichpunkt: „Forschungsdaten“)

[2] Forschungsdaten und –publikationen werden im Text zusammenfassend auch als Forschungsergebnisse bezeichnet.

[3] Mitglieder der AG Forschungsdaten: M. Kuberek (UB/HA3), A. De Nobile (ZUV/PDM) und D. Otto (Abt. V/V C 4).

[4] Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen, Forschungsdaten: www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/

Die High Level Expert Group on Scientific Data der Europäischen Kommission hat 2010 ihre Vision zu Forschungsdaten im Jahre 2030 formuliert:

> *„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."*[5]

Bereits 1998 hat die DFG auf den Wert von Forschungsdaten hingewiesen und die sichere Aufbewahrung durch die Institution, in der sie entstanden sind, in ihre Denkschrift zur „Sicherung guter wissenschaftlicher Praxis" aufgenommen:

> *„Empfehlung 7*
> *Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Institution, wo sie entstanden sind, für zehn Jahre aufbewahrt werden."*[6]

Inzwischen hat die DFG in ihren Förderrichtlinien Anforderungen zur langfristigen Sicherung und Verfügbarkeit von Forschungsdaten als Voraussetzung für die Mittelbewilligung formuliert.[7] Mit jeder Antragstellung bei der DFG verpflichtet sich der Projektleiter zur Einhaltung der Regeln der guten wissenschaftlichen Praxis.

1.1 Forschungsdaten in der TU Berlin

Auch die TU Berlin hat sich der „guten wissenschaftlicher Praxis" verpflichtet und 2002 entsprechende Richtlinien erlassen, in denen die Aufbewahrungspflicht von Primärdaten ausdrücklich aufgenommen ist.[8] Die TU Berlin ist eine traditionsreiche und renommierte Forschungsuniversität, an der fortlaufend eine Vielzahl von Forschungsprojekten durchgeführt wird, sowohl Drittmittelprojekte als auch durch die TU Berlin selbst grundfinanzierte Projekte. 2011 gab es an der TU Berlin rund 1.400 laufende Drittmittelprojekte; jährlich werden ca. 700 neue Drittmittelprojekte bewilligt.[9] Das Fördervolumen aus Drittmitteln steigt von Jahr zu Jahr kontinuierlich an; 2010 hatte es einen Umfang von über 146 Millionen Euro.[10]

Eine TU-weite Infrastruktur zur Aufbewahrung von Forschungsdaten existiert bislang nicht. Immer noch bleiben viele Daten auf den Rechnern der Wissenschaftler/innen, auf dezentralen Institutsservern oder anderen, unter Umständen unsicheren Datenträgern, und es droht die Gefahr, dass sie verlorengehen, wenn ein Projekt zu Ende ist oder Wissenschaftler/innen an eine andere Institution wechseln. In den verschiedenen Fachdisziplinen ist der Bedarf an einer institutionellen Forschungsdaten-Infrastruktur unterschiedlich ausgeprägt: So wurden in einer Reihe von Fachdisziplinen – zumeist initiiert von den jeweiligen wissenschaftlichen Fachgesellschaften – in den vergangenen Jahren weltweit verteilt disziplinspezifische Repositorien und Archive aufgebaut, an denen auch Wissenschaftler/innen der TU Berlin partizipieren. Nicht für alle Fachdisziplinen, die es an der TU Berlin gibt, existieren jedoch solche Infrastrukturen; auch bei fächerübergreifenden Forschungsvorhaben besteht Bedarf, zudem bei kleinen Projekten. Angesichts der stetig steigenden Zahl an Forschungsaktivitäten, von denen viele interdisziplinär angelegt sind, gilt es daher, in der TU Berlin eine leistungsfähige, verlässliche und nachhaltige Infrastruktur für Forschungsdaten aufzubauen. Sie umfasst neben einer technischen

[5] Riding the Wave: How Europe can gain from the rising tide of scientific data. – Final report of the High Level Expert Group on Scientific Data. A submission to the European Commission, Oktober 2010, S. 4.
cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

[6] Vorschläge zur Sicherung guter wissenschaftlicher Praxis: Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft"; Denkschrift. – Deutsche Forschungsgemeinschaft. Weinheim, 1998, Empfehlung 7, S. 12.
www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf

[7] s. Anlage: Glossar AG Forschungsdaten (D. Otto), Positionen …

[8] www.tu-berlin.de/menue/ueber_die_tu_berlin/gesetze_richt-_leitlinien/richtlinien_zur_sicherung_guter_wissenschaftlicher_praxis/

[9] Auskunft aus der Forschungsabteilung (D. Otto) im Oktober 2011

[10] www.tu-berlin.de/menue/forschung/forschungsprofil/tu-forschung_in_zahlen/

Komponente das Forschungsdatenmanagement[11], das alle Aspekte hinsichtlich der Forschungsdaten, ihrer Beschreibung, Veröffentlichung und Aufbewahrung organisiert und regelt.

1.2 Services für die TU-Wissenschaftler/innen und Mehrwert für die TU Berlin

Den Wissenschaftler/innen bringt die Infrastruktur mit ihrem umfangreichen Serviceangebot in vielerlei Hinsicht Vorteile: automatisierte Verfahren zum öffentlichen Nachweis ihrer Forschungsergebnisse (auf der eigenen Homepage, in fachspezifischen Repositorien usw.) und damit Aufwandsminimierung bei gleichzeitig erhöhter Sichtbarkeit der eigenen Forschung, eine hohe Transparenz und Vertrauenswürdigkeit der Publikationen durch Referenzierung auf supplementäre Forschungsdaten, die dauerhafte Zitierfähigkeit von Publikationen und Forschungsdaten konstante Internet-Adressen (Persistent Identifier), die nachhaltige Speicherung der Forschungsdaten und –publikationen sowie eine umfassende Beratung und Betreuung beim gesamten Prozess der Erstellung zitier- und recherchierfähiger Metadaten und der Publikation ihrer Forschungsergebnisse.

Die TU Berlin als Institution profitiert insbesondere davon, dass ihre Forschungsleistungen nach außen dokumentiert und sichtbar werden. Darüber hinaus stellt sie ihren Wissenschaftler/innen eine zukunftsgerichtete Infrastruktur und ein Dienstleistungsangebot zur Verfügung, das den Förderrichtlinien deutscher und internationaler Drittmittelgeber entspricht und ihren Ruf als renommierte Forschungsuniversität weiter fördert. Da dennoch eine Reihe von Wissenschaftler/innen dem neuen Dienstleistungsangebot zunächst zurückhaltend gegenüberstehen wird, sind in der TU Berlin Anreizsysteme zu schaffen, die die Akzeptanz erhöhen.

2 Inhalt und Rahmen des Konzeptes

Um die Begrifflichkeiten zu klären und den Rahmen für das vorliegende Konzept zum Umgang mit Forschungsdaten in der TU Berlin abzustecken, zunächst einige grundlegende Anmerkungen zu Forschungsdaten, ihrem Lebenszyklus, Repositorien und zur Aufbewahrung der Forschungsdaten.

2.1 Definition Forschungsdaten[12]

Unter Forschungsdaten sind – analog zur DFG-Definition – *„digitale und elektronisch speicherbare Daten zu verstehen, die im Zuge eines wissenschaftlichen Vorhabens z.B. durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen“*.[13]

Forschungsdaten und die zugehörigen Forschungspublikationen referenzieren gegenseitig aufeinander und werden in dem Konzept als Einheit angesehen (auch wenn es sich technisch um getrennte Datensets mit unterschiedlichen Bedingungen handelt). Beide sollen im TU-eigenen institutionellen Forschungsdatenrepositorium gleichermaßen abgelegt, verwaltet und gespeichert werden und auf den jeweils supplementären Datensatz verweisen. Ein wesentliches Kriterium und Voraussetzung für das Repositorium ist daher die technische Möglichkeit zur Referenzierung, d.h. zur Verlinkung auf einen oder mehrere supplementäre Datensätze.

Sind im vorliegenden Konzept Forschungsdaten und Forschungspublikationen gemeint, so werden sie zusammenfassend auch als **Forschungsergebnisse** bezeichnet.

[11] s. dazu: S. Büttner, H.-C. Hobohm, L. Müller: Handbuch Forschungsdatenmanagement, Potsdam 2011, opus.kobv.de/fhpotsdam/volltexte/2011/241

[12] Die Januarausgabe 2011 des D-Lib Magazine ist dem Thema „Research Data“ gewidmet und gibt eine gute Einführung in das Thema wie Zugang zu Forschungsdaten, Referenzierung von Forschungsdaten und –publikationen und eine Reihe weiterer Aspekte: www.dlib.org/dlib/january11/01contents.html

[13] s.a. Anlage: Glossar AG Forschungsdaten (D. Otto); Stichpunkt: „Forschungsdaten“

2.2 Lebenszyklus von Forschungsdaten (Research Data Lifecycle)

Die Forschungsdaten durchlaufen von ihrer Entstehung über die verschiedenen Projektphasen bis zur Archivierung und Wiederverwendung einen fortlaufenden Prozess, der in den englischsprachigen Ländern, wo man auf diesem Gebiet schon seit mehreren Jahren aktiv ist, als Research Data Lifecycle bezeichnet wird (s. Abb. 1).

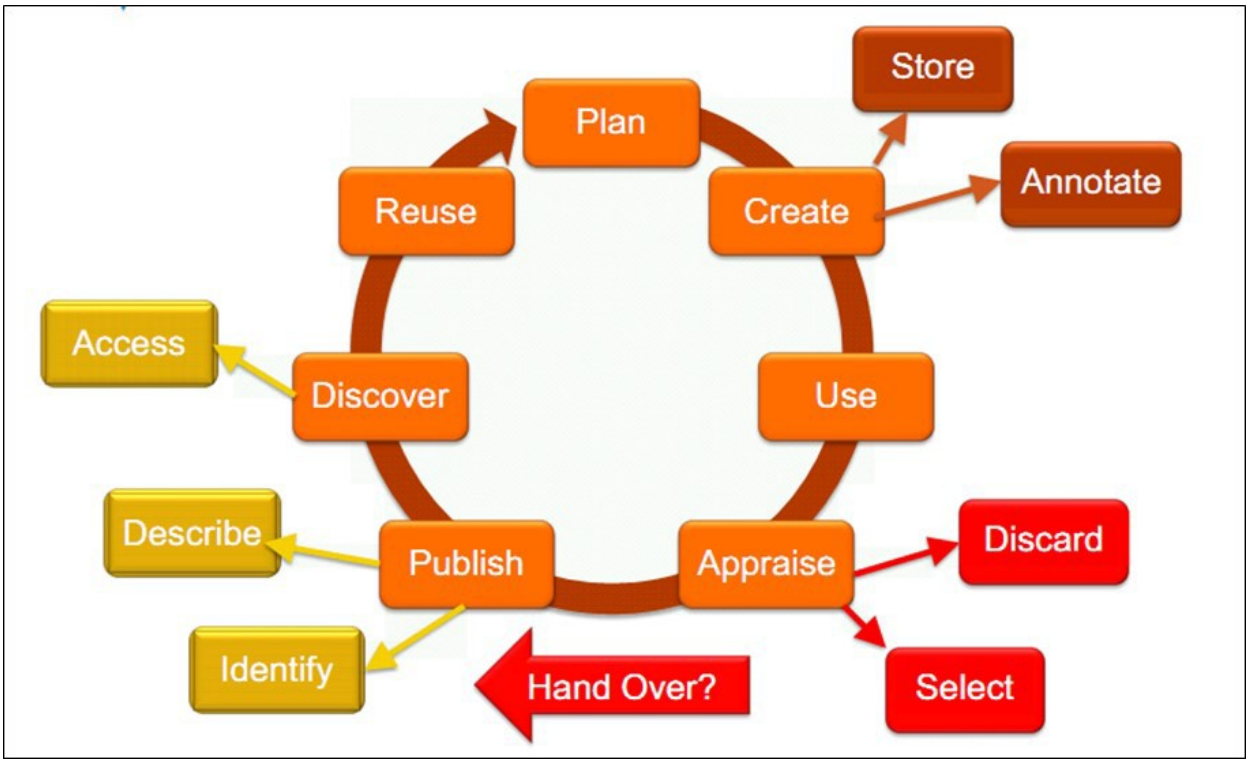


Abb. 1: Lebenszyklus von Forschungsdaten
(Quelle: S. Hodson: Supporting the Research Data Lifecycle. Folie 21 in: The JISC Managing Research Data Programme: Helping UK Universities meet research data challenges and realise benefits, 2011. www.escidoc.org/pdf/escidoc-days-2011/day1-hodson-keynote.pdf)

2.2.1 Erstellen – Aussortieren/Löschen

Werden zunächst alle möglichen Daten gesammelt, so findet im weiteren Projektverlauf ein Selektionsprozess statt: Relevante Daten werden aufbewahrt, andere aussortiert und ggf. gelöscht; dieser Prozess kann mehrere Stufen durchlaufen. Eines oder mehrere der Datensets dienen schließlich als Basis für Publikationen und werden supplementär zur jeweiligen Publikation referenziert (wobei das Forschungsdatenset seinerseits die Publikation referenziert).

2.2.2 Speichern – Archivieren

Für die Forschungsdaten, deren Ergebnisse beispielsweise in einer Publikation veröffentlicht wurden, ist es entsprechend der „guten wissenschaftlicher Praxis“ quasi ein Muss, dass sie langfristig aufbewahrt und verfügbar gemacht werden, um die Forschungsergebnisse nachvollziehen und verifizieren zu können. Eine weiteres Argument zur Aufbewahrung ist die Anforderung zur Nachnutzung der Daten: Forschungsdaten stellen eine wertvolle Quelle dar, auf der weitere Forschungen aufsetzen können.

Das Aufbewahren der Forschungsdaten ist zunächst eine Anforderung an die Wissenschaftler/innen, die nach Projektende entscheiden, welche Daten aufbewahrt werden sollen, um sie selbst weiter zu nutzen oder für die Nachnutzung öffentlich zugänglich zu machen. Das längerfristige Aufbewahren der Forschungsdaten stellt zudem bestimmte Anforderungen an das technische System, wie z.B. regelmäßige Backup-Dienste, und an die Software, wie beispielsweise Verwaltung beliebiger digitaler Objekte, Versionierung, Im- und Export und Zugriffskontrolle.

2.2.3 Metadaten – Persistent Identifier

Um die Forschungsdaten und die zugehörigen Publikationen jederzeit wiederfinden und auf sie zugreifen zu können, ist es unabdingbar, dass sie strukturiert und mit Metadaten versehen abgelegt werden.[14] Die Festlegung der Struktur und des Metadatenschemas sind daher unverzichtbare Schritte in der Vorbereitungsphase eines Forschungsprojektes, insbesondere auch, damit alle Projektmitarbeiter und –gruppen von Beginn an die erhobenen Daten einheitlich ablegen. Zur eindeutigen und dauerhaften Identifizierung und Wiederauffindbarkeit sowie zur zuverlässigen Zitierbarkeit ist jedes Datenset bzw. Subset mit einer konstanten Internet-Adresse (Persistent Identifier) zu versehen.

Es gibt eine Reihe unterschiedlicher Persistent Identifier für die Identifizierung digitaler Objekte: Der im Wissenschaftsbereich bekannteste und in den technisch-naturwissenschaftlichen Fächern am weitesten verbreitete ist DOI (Data Object Identifier).[15] Mit DataCite wurde an der TIB Hannover eine internationale DOI-Agentur für die Registrierung von Forschungsdaten aufgebaut.[16] In der deutschen Bibliothekslandschaft verbreitet ist auch URN (Uniform Resource Name), der von der Deutschen Nationalbibliothek für die zuverlässige Zitierbarkeit von Netzpublikationen eingesetzt wird; die Deutsche Nationalbibliothek betreibt einen URN-Service und bietet einen URN-Resolving-Dienst für Deutschland und die Schweiz an.[17]

2.3 Open Access

An dieser Stelle auch ein Wort zu Open Access.[18] Es ist eine grundlegende Forderung der Drittmittelgeber, dass die mit Drittmitteln erstellten Forschungsdaten Open Access zugänglich sind. Nicht alle Forschungsdaten sind jedoch für die vollständige Freigabe geeignet: Es sind z.B. Rechte von Verlagen zu beachten, manche Daten sind vertraulich, andere Daten sollen den Wissenschaftler/innen selbst als Basis für weitere Veröffentlichungen und zu weiteren Forschungen dienen usw.

Eine wesentliche Anforderung an eine geeignete Infrastruktur ist daher ein Werkzeug zur Rechteverwaltung, um sowohl den Anforderungen der Drittmittelgeber bzw. Auftraggeber als auch den Anforderungen der Wissenschaftler/innen gerecht zu werden. Während die Metadaten in der Regel frei verfügbar und recherchierbar sein sollten – und somit der **Nachweis** Open Access zugänglich ist, soll der Zugriff auf die verlinkten Datensets und Volltexte über eine Rechteverwaltung gesteuert werden: Jede/r Wissenschaftler/in soll entscheiden können, ob er den Zugriff auf die von ihm erstellten Forschungsdaten und –publikationen zum Lesen oder Download freigibt oder nicht.

2.4 Langzeitarchivierung

Die Langzeitarchivierung und Sicherstellung der Langzeitverfügbarkeit von Forschungsdaten unterliegt eigenen Bedingungen und erfordert spezielle Strategien und Maßnahmen in technischer, organisatorischer und rechtlicher Hinsicht.[19] Diese gehen über die Speicherung und Verfügbarmachung von Forschungsdaten in Repositorien weit hinaus. Für ein Archiv sind besondere Vorkehrungen zu treffen und eigene technische und organisatorische Infrastrukturen zu schaffen. Auch angesichts der hohen Kosten sind hier von der TU Berlin landesweite oder nationale Kooperationen zu prüfen und

[14] Speziell für das Publizieren und Zitieren von Forschungsdaten wurde von der DOI-Registrierungsagentur DataCite ein Metadatenschema entwickelt, an dem ein international besetztes Gremium mitgewirkt hat; Version 2.2 wurde im Juli 2011 veröffentlicht: schema.datacite.org/meta/kernel-2.2/doc/DataCite-MetadataKernel_v2.2.pdf

[15] www.doi.org

[16] www.datacite.org

[17] www.persistent-identifier.de

[18] Umfassende Informationen zu Open Access finden sich auf der Open Access Plattform: open-access.net

[19] Zur Langzeitarchivierung allgemein: www.langzeitarchivierung.de/index.htm ; das aktuell laufende Projekt LUKII (LOCKSS und kopal Infrastruktur und Interoperabilität) s. Eintrag in Wikipedia: de.wikipedia.org/wiki/LuKII ; Aktuelle Publikation: E.Euler, "Digitale Langzeitarchivierung und Distributed Storage in LuKII. Gutachten zu den Rechtsfragen, die sich im Langzeitarchivierungsprojekt LUKII (…) stellen." September 2011. edoc.hu-berlin.de/docviews/abstract.php?id=38929

anzustreben, beispielsweise mit den Berliner Hochschulen oder anderen Einrichtungen des Landes Berlin.[20]

Auch wenn Langzeitarchivierung und -verfügbarkeit hier nicht unmittelbar Thema sind, sollen einzelne Aspekte schon beim Aufbau der Forschungsdaten-Infrastruktur der TU Berlin beachtet werden. So ist unter anderem bereits beim Entstehungsprozess der Forschungsdaten darauf zu achten, dass langzeitstabile Datenformate und offene Standards verwendet, jedes digitale Objekt mit einem eindeutigen Persistent Identifier versehen, inhaltliche Metadaten strukturiert eingegeben sowie Datenformate beim Import in das Speichersystem validiert und mit notwendigen technischen Metadaten angereichert werden. Auf diese Weise erhalten die Forschungsdaten eine Qualität, wie sie für die spätere Langzeitarchivierung und -verfügbarkeit substantiell notwendig ist.

2.5 Repositorien

Repositorien hatten ursprünglich einen vor allem bibliothekarischen Fokus und wurden entwickelt, um elektronische Publikationen aufzubewahren und nachzuweisen.[21] Indem Wissenschaftler/innen die Repositorien seit einigen Jahren auch als Werkzeuge zur Ablage und Verwaltung ihrer Forschungsdaten vom Beginn des Entstehungsprozesses an nutzen, ist zu der ursprünglichen Funktion als Speicherinstrument für Publikationen eine neue Zweckbestimmung hinzugekommen. Der damit verbundene Paradigmenwechsel geht mit geänderten Anforderungen an die Repositorien-Software einher. Zu nennen sind unter anderem: Umgang mit vielen unterschiedlichen Datenformaten statt lediglich textbasierten, Verwaltung von komplexen Datensets statt einzelner Dateien, Unterstützung des Lebenszyklus von Forschungsdaten von der Entstehung der Daten bis zu ihrer Speicherung, Unterstützung kollaborativer Arbeitsprozesse. Über bisherige Repositorien-Funktionen hinausgehend erfordert der Umgang mit Forschungsdaten eine komplexe Infrastruktur, um die unterschiedlichen Anforderungen aus einzelnen Fachdisziplinen zu erfüllen. Ein weiteres wesentliches Kriterium ist die Möglichkeit zur gegenseitigen Referenzierung von Forschungsdatensets und den zugehörigen Publikationen, d.h. zur Verlinkung auf den oder die jeweils supplementären Datensätze.

Seit den 1990er Jahre ist weltweit eine große Zahl von Repositorien für elektronische Publikationen entstanden: disziplinspezifische zum Nachweis der Publikationen einzelner Wissenschafts-Communities, wie auch institutionelle zur Dokumentation der Publikationstätigkeit einzelner wissenschaftlicher Einrichtungen.[22]

Die Anfänge der Forschungsdaten-Repositorien liegen ebenfalls in den 1990er Jahren; in größerer Zahl gibt es sie seit den frühen 2000er Jahren. Eines der ersten Forschungsdaten-Repositorien in Deutschland war Mitte der 1990er Jahre *Pangaea* des Alfred-Wegener-Instituts für Klimafolgenforschung (AWI).[23] Inzwischen existieren für eine Reihe von Fach- bzw. Teildisziplinen weltweit Forschungsdaten-Infrastrukturen mit spezifischen Speichermöglichkeiten.[24] In Australien wurde in den 2000er Jahren im Rahmen einer nationalen Initiative (archer-Projekt) das Open Source

[20] So verfügt zum Beispiel das Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB) über eine prinzipiell zur Datenarchivierung geeignete Infrastruktur. 2011 ist dort das DFG-geförderte Projekt EWIG (Entwicklung von Workflowkomponenten für die Langzeitarchivierung von Forschungsdaten in den Erd- und Umweltwissenschaften) angelaufen, in dem es um die Übertragung von Forschungsdaten aus unterschiedlichen Forschungsumgebungen in digitale Langzeitarchive geht. www.geotechnologien.de/portal/gfz/cegit/Projekte/EWIG;jsessionid=91026D7C15282C895D4AB16CE07CF6F7

[21] Das älteste und umfangreichste Repositorium ist arXiv, das bereits 1991 als offen zugängliches Archiv für Preprints aus der Physik aufgesetzt wurde und heute mehr als 700.000 eDokumente aus der Physik, Mathematik, Computer Science u.a. umfasst: http://arxiv.org

[22] s. dazu beispielsweise die Auflistung von Open Access Repositorien unter oad.simmons.edu/oadwiki/Disciplinary_repositories oder OpenDOAR: www.opendoar.org
Die weltweit am häufigsten eingesetzte Repositorien-Software ist DSpace: www.dspace.org; an zweiter Stelle steht ePrints, das vor allem in Europa verbreitet ist: www.eprints.org. In Deutschland verbreitet ist zudem Opus, dessen Einsatz sich allerdings auf den deutschen Sprachraum (neben Deutschland auch mit einem kleinen Anteil in Österreich im Einsatz) beschränkt: www.kobv.de/opus4. Ein mächtiges Framework zum Aufbau von Repositorien bietet zudem Fedora Commons, www.fedora-commons.org , auf dessen Basis die Middleware eSciDoc, www.escidoc.org , aufgebaut wurde.

[23] Die beiden Repositorien des AWI, *ePIC* für Publikatonen und *Pangaea* für Forschungsdaten, bieten ein gutes Beispiel, an dem die gegenseitige Referenzierung von Publikationen und Forschungsdaten anschaulich nachvollzogen werden kann: epic.awi.de ; www.pangaea.de

[24] oad.simmons.edu/oadwiki/Data_repositories

„eResearch Toolset“ entwickelt, mit verschiedenen Softwarekomponenten zum Aufbau einer Forschungsdaten-Infrastruktur, das in den australischen Universitäten eine weite Verbreitung gefunden hat. [25] In Europa ist man von solchen nationalen Strukturen weit entfernt. In ihrem Überblick über die aktuelle Situation hinsichtlich der Forschungsdaten stellte die „Knowledge Exchange“-Kooperation, der Dänemark, Deutschland, Großbritannien und die Niederlande angehören, im November 2011 fest: „… *infrastructure is diverse, fragmented, in flux and organised differently across various disciplines in different countries*.“[26] In Europa ist Großbritannien – unterstützt und initiiert von JISC – am weitesten fortgeschritten, was den Umgang mit Forschungsdaten, Aufbau von Infrastrukturen und Archiven angeht.[27]

Zur Situation in Deutschland: Einen Überblick über Forschungsdaten-Repositorien in Deutschland gibt es nicht.[28] Die DFG fördert derzeit 27 Projekte zum Aufbau von (disziplinspezifischen) Informationsinfrastrukturen für Forschungsdaten, die im Frühjahr 2011 auf der Basis der Ausschreibung bewilligt wurden.[29] Ansonsten gibt es in Deutschland verschiedene Einzelinitiativen zum Aufbau von Forschungsdaten-Infrastrukturen. Das Land Baden Württemberg fördert mehrere Projekte zum Aufbau von Forschungsdaten-Infrastrukturen, wie z.B. „BW-eLabs“ zum „Wissensmanagement in virtuellen und remote Laboren“[30] und *„BW-eSci(T) - Entwicklung einer prototypischen e-Science-Forschungsumgebung an der Universität Tübingen“*.[31] Beide werden mittels der eSciDoc-Middleware realisiert. In den beiden anderen großen Berliner Universitäten, FU und HU, gibt es bislang keine Aktivitäten hinsichtlich ihrer Forschungsdaten und des Aufbaus eines institutionellen Repositoriums; in der HU werden erste Vorüberlegungen angestellt.

3 Organisatorisches Modell: „Servicezentrum Forschungsdaten und -publikationen“

Um die Forschungsdaten-Infrastruktur organisatorisch nachhaltig in der TU Berlin zu verankern, sieht das Konzept vor, dass UB, tubIT und Forschungsabteilung in enger Kooperation ihre Kompetenzen in einem gemeinsamen „Servicezentrum Forschungsdaten und -publikationen“ bündeln.

Ziel ist es, den Wissenschaftler/innen der TU Berlin ein dauerhaftes, zukunftsorientiertes Dienstleistungsangebot zum Umgang mit Forschungsdaten zur Verfügung zu stellen: Das Servicezentrum ist während der gesamten Dauer eines Forschungsvorhabens – von der Antragstellung über die Bereitstellung der technischen Forschungsdaten-Infrastruktur bis zu Veröffentlichung und Nachweis der Projektergebnisse – Ansprechpartner für die Wissenschaftler/innen und unterstützt sie mit umfassenden Beratungs- und Dienstleistungen. enthält unter anderem die Bereitstellung und Pflege der technischen Infrastruktur, standardisierte Metadatenschemata, Unterstützung bei Datenmanagementplänen und beim Erstellen recherchier- und zitierfähiger Metadaten, Datenpflege, Beratung beim Publizieren der Forschungsergebnisse und Open Access sowie Vermittlung grundlegender Informationen in Bezug auf Urheber- und Nutzungsrechtsfragen. Transparente

[25] archer = australian research enabling environment; archer.edu.au

[26] “A Surfboard for Riding the Wave. Towards a Four Country Action Programme on Research Data”, November 2011, S. 20. www.knowledge-exchange.info/Default.aspx?ID=469

[27] www.jisc.ac.uk/whatwedo/activitiesbytopic.aspx ; Stichpunkte auf der JISC-Homepage z.B. „Digital Repositories“, „Digital Preservation & Curation“, „Data Services & Collections“

[28] Der Aufbau eines web-basierten Nachweissystems für Forschungsdaten-Repositorien ist das Ziel des Projektes Re3Data (= Registry of Research Data Repositories). Im Rahmen des im Januar 2012 angelaufenen und bis Dezember 2013 bewilligten DFG-Projektes soll ein Nachweisinstrument für den gesamten europäischen Raum aufgebaut werden. Projektbeschreibung z.B. auf den Seiten der KIT-Bibliothek: www.bibliothek.kit.edu/cms/projekte.php

[29] www.dfg.de/foerderung/info_wissenschaft/archiv/2011/info_wissenschaft_11_18/index.html

[30] Infos zum Projekt eLabs: www.ub.uni-stuttgart.de/wirueberuns/projekte/bw-elabs ; Website: www.bw-elabs.org ; M. Razum: BW-eLabs: “Data acquisitions in the Lab”. Oktober 2011: eSciDoc-Days 2011: www.escidoc.org/pdf/escidoc-days-2011/day2-razum-bwelabs.pdf ; Abschlussveranstaltung im Januar 2012: www.bw-elabs.org/news/news_0003.html?_locale=de

[31] Infos zum Projekt: eSci(T): www.uni-tuebingen.de/einrichtungen/zentrum-fuer-datenverarbeitung/projekte/bw-escit/projektuebersicht/escidoc.html

Kommunikations- und Informationsstrukturen sorgen dafür, dass Wissenschaftler/innen in jeder Phase eines Forschungsprojektes wissen, wohin sie sich wenden können.

Eine Informationsplattform im Internet gibt den Wissenschaftler/innen weitere Hilfsmittel zum Umgang mit Forschungsdaten an die Hand, wie beispielsweise Checklisten, Zusammenstellung und Links zu disziplinspezifischen Repositorien und Datencentern usw. Beim Aufbau dieser Plattform kann auf eine Fülle bereits existierender Materialien, die den gesamten Lebenszyklus der Forschungsdaten umfassen, zurückgegriffen werden.[32] So gibt es solche Plattformen zum Forschungsdatenmanagement mit sehr umfangreichen Informationsmaterialien und Schulungsunterlagen beispielsweise bei den Universitäten von Cambridge,[33] Glasgow[34] und Edinburgh[35] wie auch beim MIT (Massachusetts Institute of Technology).[36] Die überwiegende Mehrzahl der Materialien steht unter der Creative Commons Licence im Internet zur freien Nachnutzung zur Verfügung. Für den Aufbau des TU-Informationsangebotes gilt es, die bereits vorhandenen Materialien zu sichten und ggf. in Zusammenarbeit mit den Wissenschaftler/innen an den konkreten Bedarf anzupassen. Die Zielsetzung ist es, den Wissenschaftler/innen ein Portfolio an die Hand zu geben, das zur schnellen Erstinformation im Internet und auch zum Nachschlagen geeignet ist.

Auch ein Schulungsangebot zum Forschungsdatenmanagement, beispielsweise zu Formaten, Metadaten und ihrer strukturierten Ablage, soll ins Auge gefasst werden. Hier gilt es, zunächst den potenziellen Bedarf zu ermitteln und entsprechende Schulungsunterlagen zu entwickeln.

3.1 Nutzung vorhandener Kompetenzen

In der TU Berlin sind in den drei Infrastruktureinrichtungen UB, tubIT und der Forschungsabteilung spezifisches Know-how und einschlägige Kompetenzen vorhanden, auf die beim Aufbau und Betrieb des Servicezentrums zurückgegriffen werden kann. Insbesondere der UB als Informationsinfrastruktureinrichtung, zu deren Kernaufgaben das Angebot von Informations- und Beratungsdiensten, der Umgang mit Metadaten sowie die Bereitstellung von Nachweis- und Recherchesystemen gehören, kommt beim Umgang mit Forschungsdaten eine zentrale Rolle zu.

Bezogen auf das Forschungsdatenmanagement, verfügen UB, tubIT und Forschungsabteilung u.a. über folgende Kompetenzen:

- UB:
  Umgang mit Metadaten-Erstellung und Metadaten-Formate, formale Qualitätssicherung von Daten, Datenpflege, Persistent Identifier, Nachweis und Recherche, Verfahren und Workflows für elektronisches Publizieren, Lizenzierungsfragen, Urheberrecht, Open Access, …
- tubIT:
  Installation und Pflege von Hard- und Software zum Aufbau und Betreiben der technischen Infrastruktur, Backup-Services, Authentizifierung/Autorisierung (Shibboleth), Standardprotokolle und –schnittstellen, Bereitstellung von Werkzeugen für kollaborative Arbeitsweisen, …
- Forschungsabteilung:
  Informationen über Projekte, Drittmittel und ihre Verwaltung, Projektanträge, Informationen über Drittmittelgeber und deren Anforderungen, …

Die folgende Skizze verdeutlicht, ausgehend von den grundlegenden Bereichen des Forschungsdatenmanagements, in welchen Segmenten und auf welche Weise die genannten Kompetenzen zum Einsatz kommen können (s. Abb. 2). Auch den Wissenschaftler/innen kommt beim Forschungsdatenmanagement eine wesentliche Rolle zu: Sie sind es, die die Forschungsdaten erzeugen und sie daher inhaltlich kennen und beurteilen können. Auch ihre Rolle wird in der Skizze deutlich.

[32] Hier nimmt Großbritannien, wo JISC (Joint Information Systems Committee) in dieser Hinsicht schon seit Jahren große Aktivitäten entfaltet, die führende Rolle in Europa und eine der führenden Rollen weltweit ein. www.jisc.ac.uk

[33] www.lib.cam.ac.uk/dataman

[34] www.gla.ac.uk/services/datamanagement

[35] www.ed.ac.uk/schools-departments/information-services/services/research-support/data-library/research-data-mgmt

[36] libraries.mit.edu/guides/subjects/data-management

Abb. 2: Einsatz der vorhandenen Kompetenzen beim Forschungsdatenmanagement

3.2 Verteilte Zuständigkeiten und strukturierte Workflows

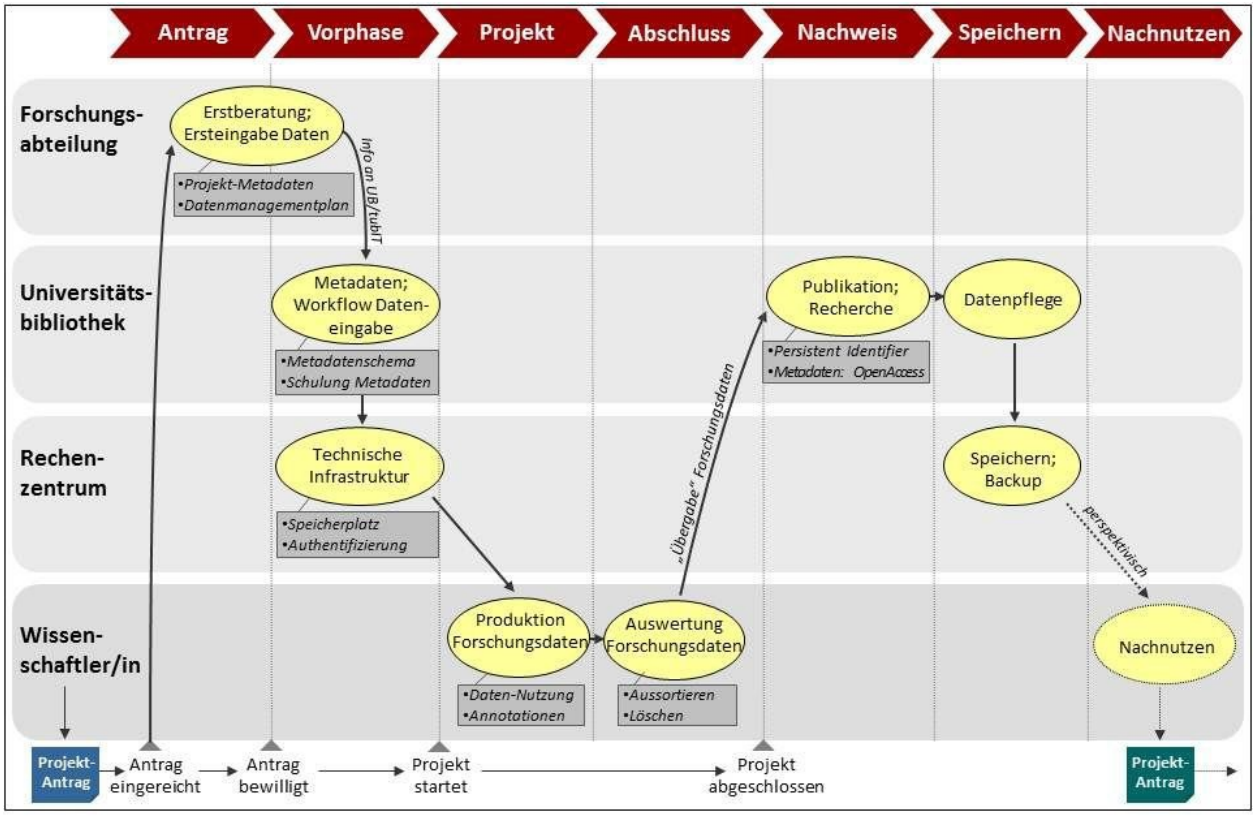
Um die Zuverlässigkeit und Dauerhaftigkeit des Dienstleistungsangebotes zu gewährleisten, wird die Arbeitsteilung zwischen den drei beteiligten Einrichtungen verbindlich geregelt. Dazu werden – dem Forschungsdaten-Lebenszyklus[37] entsprechend (s. Abb. 1) – für jede Phase eines Forschungsprojektes die Zuständigkeiten hinsichtlich der Forschungsdaten und strukturierte Workflows zwischen den drei beteiligten Infrastruktureinrichtungen definiert (s. Abb. 3):

- Die Forschungsabteilung, die als erste der drei Einrichtungen von einem Projektvorhaben erfährt, führt die Erstberatung durch. Sie liefert in einem automatisierten Verfahren erste formale Informationen (Titel, Wissenschaftler/in, …) in den Projekt-Metadatensatz und beginnt den Datenmanagementplan, eine strukturierte Checkliste,[38] in der sie Angaben zum Drittmittelgeber, seinen Anforderungen und weitere allgemeine Informationen zum Projekt dokumentiert. Sobald die Bewilligung vorliegt, werden von UB und tubIT vorbereitende Arbeiten für das Forschungsprojekt aufgenommen. Der gesamte Komplex Metadaten liegt – auch im weiteren Verlauf des Forschungsdatenmanagements – in der Zuständigkeit der UB. In der Phase der

[37] Siehe z.B. S. Hodson: *The JISC Managing Research Data Programme: Helping UK Universities meet research data challenges and realise benefits*, 2011, Folie 21: Supporting the Research Data Lifecycle. www.escidoc.org/pdf/escidoc-days-2011/day1-hodson-keynote.pdf

[38] Siehe dazu beispielsweise die im Dezember 2011 vom BMBF-geförderten Projekt WissGrid (www.wissgrid.de/index.html) erstmals für den deutschsprachigen Raum veröffentlichte Checkliste und Leitfaden zum Forschungsdatenmanagement:
Checkliste: www.wissgrid.de/publikationen/deliverables/wp3/WissGrid-oeffentlicher-Entwurf-Checkliste-Forschungsdaten-Management.pdf
Leitfaden: www.wissgrid.de/publikationen/deliverables/wp3/WissGrid-oeffentlicher-Entwurf-Leitfaden-Forschungsdaten-Management.pdf
(Einen guten Überblick über die Aktivitäten der D-Grid Initiative gibt die Broschüre „D-Grid – Die Deutsche Grid-Initiative: Vorstellung der Projekte“ [leider ohne Jahr, 2009?]: www.d-grid-ggmbh.de/downloads/BroschuereAHM2009.pdf)

Projektvorbereitung geht es dabei u.a. um den Workflow für die Dateneingabe, der – ebenso wie das Metadatenschema – eventuell an die fachspezifischen Projekterfordernisse anzupassen ist, und die Schulung der Wissenschaftler/innen hinsichtlich der Erstellung strukturierter, recherchefähiger Metadaten. tubIT ist für die technische Infrastruktur zuständig, stellt u.a. den benötigten Speicherplatz zur Verfügung und passt die Rechteverwaltung an, entsprechend den Projektanforderungen.



„Forschungsdaten-Lifecycle" – Zuständigkeiten und Workflows © Monika Kuberek, Januar 2012

Abb. 3: Verteilte Zuständigkeiten und strukturierte Workflows im „Servicezentrum Forschungsdaten und -publikationen" der TU Berlin

- Auch den Wissenschaftler/innen kommt beim Forschungsdatenmanagement eine wesentliche Rolle zu, da sie die Forschungsdaten produzieren, diese inhaltlich kennen und sie beurteilen können. Inhaltliche Informationen zu den Daten, wie Annotationen, geben die Wissenschaftler/innen daher selbst in den Projekt-Metadatensatz ein. Vor dem Abschluss des Projektes begutachten sie die Daten und werten sie aus – auch im Hinblick auf eine eventuelle spätere Wiederverwendung. Dabei wählen sie die Forschungsdaten aus, die aufbewahrt werden sollen, und löschen die Forschungsdaten, die nicht mehr benötigt werden.
- Nach Abschluss des Projektes werden die Forschungsergebnisse an die UB „übergeben": Sie unterstützt den öffentlichen Nachweis der in der TU Berlin erbrachten Forschungsleistungen durch elektronisches Publizieren der Forschungsergebnisse, sorgt unter anderem dafür, dass die Datensätze mit Persistent Identifier versehen und damit dauerhaft zitierfähig werden und berät die Wissenschaftler/innen hinsichtlich Open Access und vermittelt grundlegende Informationen zu Urheber- und Nutzungsrechtsfragen. Mit dem Wissensportal Primo betreibt die UB ein Rechercheinstrument, mit dem die Forschungsergebnisse im Internet weltweit nachgewiesen, suchbar und – falls lizenzrechtlich erlaubt – zugreifbar sind. Im weiteren Forschungsdaten-Lebenszyklus ist die UB für die Datenpflege zuständig, tubIT für das Speichern und regelmäßige Sichern der Daten.

Während der gesamten Projektphase ist das Servicezentrum direkter Ansprechpartner für die Wissenschaftler/innen und sorgt dafür, dass Fragen schnell und kompetent beantwortet werden.

3.3 Organisationsstruktur

Wesentlich für das Servicezentrum ist eine schlanke, effiziente Organisationsstruktur. Ein kleines Steuerungsgremium, dem jeweils eine Person aus der UB, tubIT und der Forschungsabteilung angehören, lenkt und koordiniert die Arbeiten und ist für die Weiterentwicklung der Forschungsdatenkonzepte in der TU Berlin zuständig. Die Leitung liegt bei der UB, deren Kerngeschäft als Informationsinfrastruktur-Einrichtung der TU Berlin eine hohe Affinität zum Themengebiet Forschungsdaten und deren Management aufweist.

4 Technische Infrastruktur

Vor dem Hintergrund des großen Spektrums unterschiedlicher Fachdisziplinen muss eine für den Einsatz in der TU Berlin geeignete Infrastruktur in der Lage sein, mit Forschungsdaten vieler Disziplinen umzugehen. Gleichzeitig muss das Forschungsdaten-Repositorium die Möglichkeit zur Speicherung und gegenseitigen Referenzierung von Forschungsdaten und -publikationen bieten.[39]

Unter Berücksichtigung dieser Aspekte kommen für den Einsatz in der TU Berlin prinzipiell die Software-Systeme DSpace, EPrints und eSciDoc in Frage. Alle drei Software-Systeme wurden mit der Zielsetzung der Verwaltung, Aufbewahrung und Nachnutzung von Forschungsdaten und -publikationen entwickelt und werden weltweit für institutionelle wie auch disziplinspezifische Repositorien eingesetzt. Sie stehen als freie Open-Source-Software zur Verfügung.

- DSpace[40] ist mit über 1.000 Installationen die weltweit am weitesten verbreitete Repositorien-Software. Die Software wurde vom Massachusetts Institute of Technology (MIT) gemeinsam mit HP Labs (Hewlett Packard) in Cambridge, Massachusetts, entwickelt; 2002 wurde die erste Version freigegeben. Seit 2009 stehen das DSpace Project und die DSpace Community unter der Verantwortung der Non Profit Organisation DuraSpace. Die Weiterentwicklung erfolgt durch eine internationale Entwickler-Community.
- EPrints[41] steht in der weltweiten Verbreitung an zweiter Stelle; die Software wird insbesondere in Europa eingesetzt. Sie wurde an der University of Southampton entwickelt und erstmals im Jahre 2000 freigegeben. Die Weiterentwicklung liegt bei der University of Southampton, unterstützt von einer internationalen Entwickler-Community. Erweiterungen und Verbesserungen der Software werden kontinuierlich vorangetrieben; Release 3.3.8 wurde im Januar 2012 veröffentlicht.
- eSciDoc[42] ist eine relativ neue Entwicklung. Die auf Fedora Commons aufbauende Middleware wurde vom FIZ Karlsruhe und der Max-Planck-Digital Library (MPDL) im Rahmen eines BMBF-geförderten Projektes entwickelt und 2009 in einer ersten Version freigegeben. Eingesetzt wird die Software in Deutschland beispielsweise in der Max-Planck-Gesellschaft und im GeoForschungsZentrum Potsdam, zudem international in verschiedenen Institutionen unter anderem in Dänemark, Japan, Spanien. FIZ Karlsruhe und MPDL haben ein Memorandum zur gemeinsamen Weiterentwicklung von eSciDoc unterzeichnet; eine internationale Entwicklergemeinde ist im Aufbau begriffen.

Die Repository-Software muss flexibel für verschiedene Anforderungen anpassbar sein und geeignet, die Anforderungen an ein Forschungsdaten-Repositorium für die TU Berlin zu erfüllen:

[39] Ausführliche Informationen zur organisatorischen und technischen Planung von Repositorien, Checklisten usw. sind auf den Seiten von JISC zusammengestellt: http://www.jiscinfonet.ac.uk/infokits/repositories

[40] www.dspace.org

[41] www.eprints.org

[42] www.escidoc.org und www.fiz-karlsruhe.de/escidoc_project.html?&L=1

- Bereitstellung einer disziplinunabhängigen Infrastruktur zum Einsatz in einer multidisziplinären Einrichtung; softwareunabhängiges Speicherformat, offene Standardschnittstellen.
- Verwaltung und Speicherung beliebiger Arten von digitalen Objekten (flexible Content-Modelle, beliebige Datenformate und Metadaten-Profile, …).
- Gewährleistung einer dauerhaften und vertrauenswürdigen Zitierfähigkeit der Forschungsergebnisse durch Unterstützung einer Vielzahl von Persistent Identifiern.
- Ein fein granuliertes Rechtemanagement und damit die Wahlmöglichkeit, die Forschungsergebnisse (Datensets und Volltexte) frei zugänglich im Internet anzubieten (Open Access) oder lediglich die Recherche der Metadaten zuzulassen, nicht aber den Zugriff auf die Forschungsergebnisse selbst.

Gleichzeitig muss die Software Funktionen bereitstellen, um die Vertrauenswürdigkeit von Forschungsergebnissen zu fördern: So müssen traditionelle Publikationen auf Forschungsdaten (als supplementäre Materialien) referenzieren und die Ergebnisse somit verifiziert werden können. Unterstützt werden muss auch die nachhaltige Speicherung und Wiederverwendbarkeit der Forschungsdaten: Beschreibung und Kontext von Forschungsdaten müssen dokumentiert und gespeichert werden können, eine wesentliche Voraussetzung für die Nachnutzung und Archivierung von Forschungsdaten.

5 Umsetzung des Konzepts: Projekt „Forschungsdaten“

Der Aufbau der Forschungsdaten-Infrastruktur soll im Rahmen eines zweijährigen Projektes erfolgen. Das Projekt soll in der ersten Jahreshälfte 2012 beginnen.

- Die Arbeitsgruppe wird das Konzept im Frühjahr 2012 in einem kleinen Workshop innerhalb der TU Berlin vorstellen und diskutieren. Zu dem Workshop sollen insbesondere auch – als potenzielle Pilotpartner – interessierte Wissenschaftler/innen eingeladen werden.
- Die Umsetzung des organisatorisch-technischen Konzeptes erfolgt in enger Zusammenarbeit mit einem Pilotpartner. Die AG Forschungsdaten, die das vorliegende Konzept erstellt hat, wird das Projekt begleiten; die Federführung liegt bei der UB.
    - Im Rahmen des Projektes wird das „Servicezentrum Forschungsdaten und -publikationen“ eingerichtet und die Kooperation zwischen UB, tubIT und Forschungsabteilung etabliert. Dazu werden die künftigen Kommunikationsstrukturen aufgebaut, die vorgesehene Aufgabenteilung und die Workflows verifiziert und im Detail ausgearbeitet.
    - Während des Projektes wird mit dem Aufbau der Internet-Informationsplattform begonnen. Dabei kann auf bereits erarbeitete Materialien, insbesondere aus Großbritannien und den USA, zurückgegriffen werden, die unter der Creative Commons Licence im Internet zur freien Nachnutzung zur Verfügung stehen. Die Materialien gilt es zu sichten und ggf. in Zusammenarbeit mit den Wissenschaftler/innen an den konkreten Bedarf anzupassen.

Mit den Vorarbeiten – Auswahl einer Repository-Software, Aufbau einer Testumgebung, Testinstallation – soll nach dem Workshop möglichst umgehend begonnen werden. Für die Installation und erste Implementierungsarbeiten werden 2-3 Monate veranschlagt. Die Anpassung der Software erfolgt in Zusammenarbeit mit dem Pilotpartner.

6 Perspektive: Gesamtnachweis der in der TU Berlin erbrachten Forschungsleistungen

6.1 Nachweis und Recherche der Forschungsergebnisse

Über Standardprotokolle soll die ausgewählte Repository-Software einen umfassenden Nachweis der Forschungsergebnisse im Internet ermöglichen, der die Forschungsleistung der einzelnen TU-Wissenschaftler/innen und der TU Berlin als Institution nach außen dokumentiert und sichtbar macht. Dazu werden Verzeichnisse automatisiert aufgebaut und laufend ergänzt, so beispielsweise

- der Nachweis der wissenschaftlichen Tätigkeit auf der Homepage einer Wissenschaftlerin oder eines Wissenschaftlers,
- der Nachweis in den Web-Verzeichnissen der jeweiligen wissenschaftlichen Community,
- eine Hochschulbibliographie als öffentliches Nachweisinstrument der an der TU Berlin entstandenen Forschungsergebnisse,
- ein Preprint-Server zur Vorab-Veröffentlichung von Forschungsergebnissen,
- elektronische Publikationen mit referenzierten Forschungsdaten.

Die Recherche nach den Forschungsergebnissen wird über das Wissensportal Primo der UB realisiert. Primo ist ein "One-Stop-Portal" mit einer Suchmaschine, die gleichzeitig sämtliche Ressourcen der UB durchsucht: Online-Katalog, Dokumentenserver, Nationallizenzen, lizenzierte elektronische Zeitschriften und E-Books sowie zahlreiche freie und lizenzierte Datenbanken – und künftig auch das institutionelle Forschungsdaten-Repositorium.

6.2 Perspektive Gesamtnachweis

In einer Reihe von Fachdisziplinen wurden in den vergangenen Jahren weltweit verteilt disziplinspezifische Infrastrukturen für Forschungsdaten aufgebaut. Auf diese Weise sind einzigartige Wissensspeicher entstanden, an denen auch Wissenschaftler/innen der TU Berlin ihren Anteil haben. Die TU-Wissenschaftler/innen sollen auch künftig die für ihre Disziplin bereits existierenden Forschungsdaten-Infrastrukturen so weit wie möglich nutzen und sich am Auf- und Ausbau der disziplinspezifischen globalen Forschungsdatenspeicher beteiligen. Der starken Verankerung der Wissenschaftler/innen in ihren Fachdisziplinen trägt das vorliegende Konzept Rechnung, indem es eine verteilte Datenhaltung für die von TU-Wissenschaftler/innen erbrachten Forschungsleistungen vorsieht: im institutionellen Forschungsdaten-Repositorium der TU Berlin und in den weltweit verteilten disziplinspezifischen Repositorien.

Um die gesamte in der TU Berlin erbrachte Forschungsleistung zu dokumentieren, ist perspektivisch ein Gesamtnachweis – inkl. Link zu den Datensets und Volltexten – aller von Wissenschaftler/innen der TU Berlin erstellten Forschungsdaten und der zugehörigen Publikationen grundsätzlich möglich (s. Abb. 4). Mittels automatisierter Verfahrens über offene Standardprotokolle wie OAI-PMH und SWORD können die Metadaten inkl. Persistent Identifier eingesammelt, im TU-Forschungsdaten-Repositorium gespeichert und über das Wissensportal Primo suchbar gemacht werden. Mit Hilfe von Primo kann die UB den zentralen Einstieg in die Metadaten-Recherche aller Forschungsergebnisse der TU Berlin organisieren. Über konstante Internet-Adressen (Persistent Identifier) sind darüber hinaus die entfernt liegenden Volltexte bzw. Forschungsdatensets prinzipiell zugreifbar: Entsprechend der Entscheidung der Wissenschaftler/innen können sie gesperrt, nur für bestimmte Gruppen oder frei im Netz zugänglich gemacht werden.

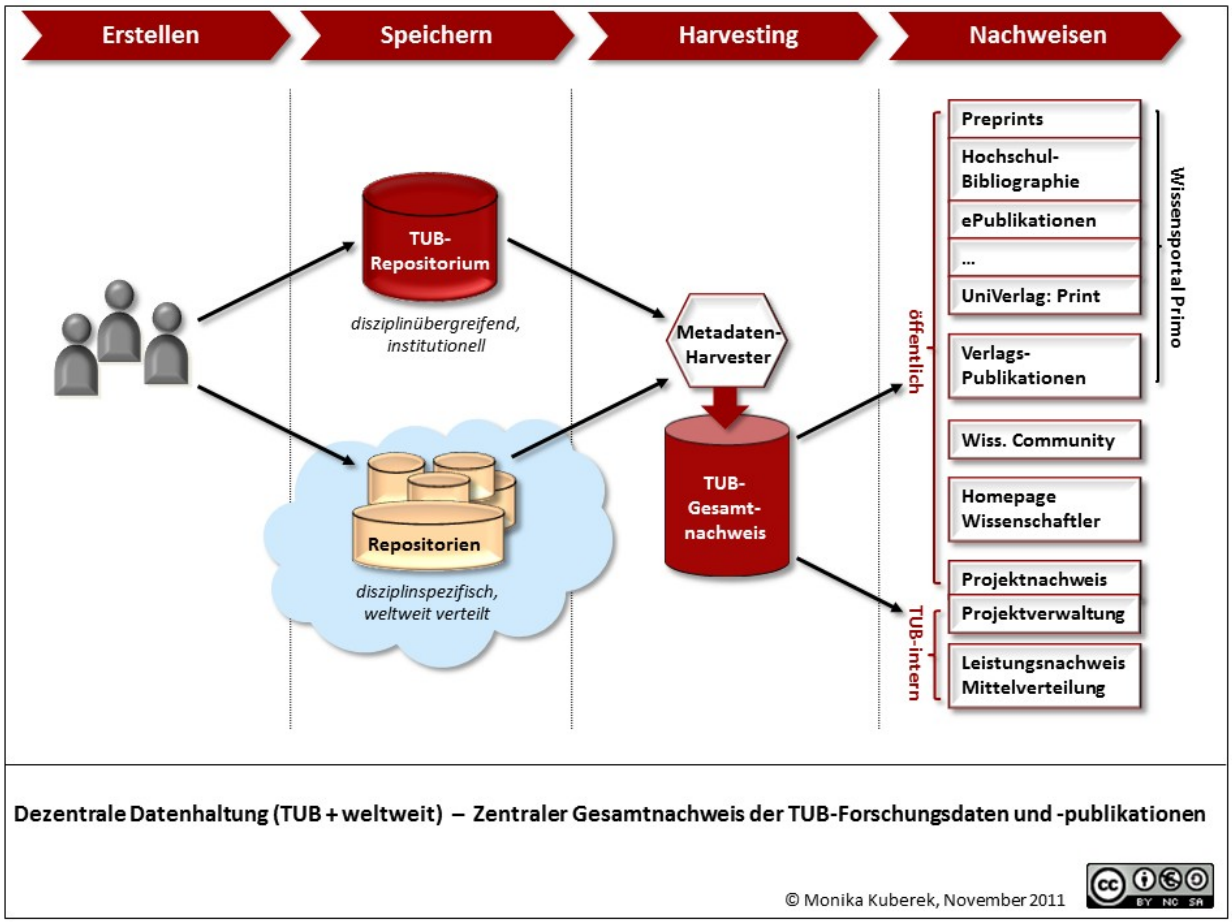


Abb. 4: Perspektive Gesamtnachweis der in der TUB entstandenen Forschungsdaten und -publikationen

Perspektivisch könnten auch die nicht-öffentlichen Nachweisinstrumente der TU Berlin davon profitieren, dass Metadaten zu Forschungsdaten und –publikationen nur einmal erfasst und automatisiert übernommen werden:

- das Verzeichnis für den Leistungsnachweis und die Mittelvergabe von LinF (Leistungserfassung in Forschung und Lehre)
- die Verzeichnisse für die Projektverwaltung der Forschungsabteilung

Voraussetzung dafür ist, dass die für LinF und die Forschungsabteilung eingesetzten Software-Systeme Standardschnittstellen zum Austausch der Metadaten, wie beispielsweise OAI-PMH, bedienen können.

Anlage: Glossar AG Forschungsdaten (D. Otto)

Version 1: 03.01.2012 (erstellt durch D. Otto)
Version 2: 14.05.2012 (aktualisiert und ergänzt durch D. Otto)
Version 3: 08.08.2012 (aktualisiert und ergänzt durch S. Aust)

Glossar AG Forschungsdaten[43]

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Positionen zum Umgang mit Forschungsdaten und Forschungsergebnissen

A	
Allianz der deutschen Wissenschafts-organisationen	siehe Positionen
Archivserver (engl.: repository)	„Als Archivserver wird eine technische und organisatorische Infrastruktur bezeichnet, deren Aufgabe vorrangig die Speicherung, Verwaltung und Langzeitarchivierung digitaler Objekte ist. Systeme des Typs 'Archivserver' werden im Kontext der Langzeitarchivierung getrennt von Systemen des Typs Dokumentenserver betrachtet, um den jeweiligen unterschiedlichen Schwerpunkten der Funktionalität gerecht zu werden.“ ***nestor-Glossar: Archivserver*** http://lza.ddb.de/index.php?module=Encyclopedia&func=displayterm&id=52&vid=1
B	
BMBF	siehe Positionen
C	
CODATA	“Das Committee on Data for Science and Technology ist eine […] Organisation mit dem Ziel der Verbesserung von Qualität, Zuverlässigkeit und Zugänglichkeit von interessanten Daten aus allen Feldern der Wissenschaft und Technologie. CODATA wurde um 1966 vom Internationalen Wissenschaftsrat (*International Council for Science*) gegründet.“ ***Wikipedia: Committee on Data for Science and Technology*** http://de.wikipedia.org/wiki/CODATA
D	
Datenmanage-ment	„das Datenmanagement ist die Menge aller methodischen, konzeptionellen, organisatorischen und technischen Maßnahmen und Verfahren zur Behandlung von ‚Daten’ [mit dem Ziel] der Gewährleistung einer maximalen Unterstützung auszuführender Geschäftsprozesse …“ ***Bettina Berendt , Joaquin Vanschoren, Bo Gao: Datenanalyse und -visualisierung, in: Handbuch Forschungsdatenmanagement, hrsg. v. Stephan Büttner, Hans-Christoph Hobohm, Lars Müller, Bad Honneff, 2011, S. 140.*** http://opus.kobv.de/fhpotsdam/volltexte/2011/241/pdf/HandbuchForschungsdatenmanagement.pdf
Datenmanage-mentplan (DMP)	unterstützende Tools und Guidelines zur Erstellung eines Datenmanagementplans: • University of California Curation Center of the California Digital Library: Guidance and Resources you’re your Data Management Plan: https://dmp.cdlib.org/ (user: otto passwort: tu-berlin1)

[43] Wenn kein anderes Datum vermerkt, letzter Zugriff auf die Internetseiten am 14.05.2012.

Am 08.08.2012 wurde von S. Aust (Referendarin UB TU Berlin) die Aktualität der Links überprüft, falls notwendig aktualisiert und mit dem aktuellen Zugriffsdatum 08.08.2012 versehen. Zudem wurden die neueren Dokumente von DFG, Wissenschaftsrat und EU eingepflegt.

	• DCC - Digital Curation Centre, UK: Data Management Planning Tool: https://dmponline.dcc.ac.uk/ (user: otto passwort: tu-berlin) • archer = australian research enabling environment: eResearch Toolset: http://archer.edu.au/) • Australian National Data Service http://ands.org.au/resource/data-management-planning.html • MIT: Data Management and Publishing http://libraries.mit.edu/guides/subjects/data-management/ • Cornell University: Data Management Planning Guide https://confluence.cornell.edu/display/rdmsgweb/data-management-planning-overview • Hinweise auf laufende Aktivitäten zur Entwicklung von DPM und DPM-Tools auch bei JISC: http://www.jisc.ac.uk/whatwedo/activitiesbytopic.aspx
DFG	siehe Positionen
digital preservation	s. Langzeitarchivierung
DINI	siehe Positionen
Dokumenten-server	Server, auf dem Dokumente wie z.B. Forschungsberichte oder Projektergebnisse gespeichert werden können.
E	
EU	siehe Positionen
EUROHORCs - European Heads of Research Councils / **ESF** – European Science Foundation	siehe Positionen
F	
Forschungs-daten	=Rohdaten , =Primärdaten „Unter Forschungsdaten sind im Sinne dieser Fördermaßnahme digitale und elektronisch speicherbare, Daten zu verstehen, die im Zuge eines wissenschaftlichen Vorhabens z.B. durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.“ ***DFG: Ausschreibung „Informationsinfrastrukturen für Forschungsdaten“, S.1*** http://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_forschungsdaten_1001.pdf „Dieser Begriff „Primärdaten“ sorgt immer wieder für Diskussion, denn die Definition des Begriffs ist sehr von der eigenen Rolle in der wissenschaftlichen Wertschöpfungskette bestimmt. […] Der begrifflichen Klarheit wegen sollte daher das Präfix „Primär-“ nicht mehr verwendet werden und statt dessen nur noch von wissenschaftlichen Daten oder Forschungsdaten gesprochen werden.“ ***Jens Klump: Digitale Forschungsdaten, in: nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, hg. v. H. Neuroth, A. Oßwald, R. Scheffel, et al., Göttingen, 2010 [Version 2.3], S. 104-105.*** http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf **„Was sind Forschungsdaten?** Die verschiedenen Wissenschaften sprechen von Messdaten, Rohdaten, empirischen Daten, Quelldaten, Forschungsrohdaten etc., verwenden viele unterschiedliche Begriffe und haben auch unterschiedliche Konzepte von dem, was für sie ein „Datum“ ist. Mittlerweile kristallisiert sich heraus, dass im deutschen Sprachgebrauch der Begriff

	„Forschungsdaten“ bevorzugt wird.“ ***Stephan Büttner, Hans-Christoph Hobohm, Lars Müller: Research Data Management, in: Handbuch Forschungsdatenmanagement, hrsg. v. Stephan Büttner, Hans-Christoph Hobohm, Lars Müller, Bad Honneff, 2011, S. 15.*** http://opus.kobv.de/fhpotsdam/volltexte/2011/241/pdf/HandbuchForschungsdatenmanagement.pdf
Forschungs-projektdaten	Metadaten zu Forschungsprojekten
Forschungs-publikation	Publikation von Forschungsergebnissen in einem Journal oder einem Dokumentenserver, die ein peer reviewed Verfahren durchlaufen haben
L	
Langzeit-archivierung (engl.: digital preservation, long-term preservation)	„‚Langzeit‘ ist die Umschreibung eines nicht näher fixierten Zeitraumes, währenddessen wesentliche nicht vorhersehbare technologische und soziokulturelle Veränderungen eintreten, die sowohl die Gestalt als auch die Nutzungssituation digitaler Ressourcen in rasanten Entwicklungszyklen vollständig umwälzen werden.“ ***Hans Liegmann: Einführung, in: nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, hg. v. H. Neuroth, A. Oßwald, R. Scheffel, et al., Göttingen, 2010 [Version 2.3], S.3*** http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf
Langzeit-verfügbarkeit (engl.: long-term availability)	„Erhaltung der Langzeitverfügbarkeit digitaler Objekte umfasst alle Maßnahmen, die dazu dienen, digitale Objekte für unsere „Nachwelt“ benutzbar zu erhalten. Mit „unserer Nachwelt“ ist eine unbegrenzt in die Zukunft reichende Zeitspanne gemeint.“ ***nestor-Glossar: Langzeitverfügbarkeit*** http://lza.ddb.de/index.php?module=Encyclopedia&func=displayterm&id=7&vid=1
long-term availability	s. Langzeitverfügbarkeit
long-term preservation	s. Langzeitarchivierung
N	
nestor	Kompetenznetzwerk zur digitalen Langzeitarchivierung: http://www.langzeitarchivierung.de/
O	
OAI	„Open Archives Initiative: Initiative von Betreibern von Preprint-Servern und anderen Dokumentenservern, um die auf diesen Servern abgelegten elektronischen Publikationen im Internet besser auffindbar und nutzbar zu machen.“ ***Wikipedia: Open Archives Initiative*** http://de.wikipedia.org/wiki/Open_Archives_Initiative
OAI PMH	„OAI Protocol for Metadata Harvesting (OAI-PMH) zum Einsammeln und Weiterverarbeiten von Metadaten“ ***Wikipedia: Open Archives Initiative*** http://de.wikipedia.org/wiki/Open_Archives_Initiative
OAIS (engl.: Open Archival Information System)	„Das als ISO 14721 verabschiedete Referenzmodell beschreibt ein Archiv als Organisation. In dieser wirken Menschen und Systeme mit der Aufgabenstellung zusammen, Informationen zu erhalten und einer definierten Nutzerschaft verfügbar zu machen. OAIS-Konformität beinhaltet die Übernahme der Verantwortung für die Einhaltung der im Referenzmodell vorgegebenen Rahmenbedingungen. Das Referenzmodell spezifiziert weder das Design noch die Implementierung eines OAIS-konformen Archivs.“ ***kopal: Standards: OAIS (Open Archival Information System)*** http://kopal.langzeitarchivierung.de/index_standards.php.de
OECD	siehe Positionen
OpenAIRE	„OpenAIRE aims to support the implementation of Open Access in Europe. It provides the means to promote and realize the widespread adoption of the **Open Access Policy**, as set out by the **ERC Scientific Council Guidelines for Open Access** and the **Open**

	Access pilot launched by the European Commission. OpenAIRE, a three-year project, will establish the infrastructure for researchers to support them in complying with the EC OA pilot and the ERC Guidelines on Open Access. It will provide an extensive European Helpdesk System, based on a distributed network of national and regional liaison offices in 27 countries, to ensure localized help to researchers within their own context. It will build an OpenAIRE portal and e-Infrastructure for the repository networks and explore scientific data management services together with 5 disciplinary communities. It will also provide a repository facility for researchers who do not have access to an institutional or discipline-specific repository." ***OpenAire: Fact Sheet*** http://www.openaire.eu/en/about-openaire/general-information/fact-sheet
OpenAIREplus	„OpenAIREplus erweitert das Ziel der Open Access-Verfügbarkeit von Forschungspublikationen, indem es wechselseitige Verknüpfungen von Publikationen zu dazugehörigen Forschungsdaten und Förderformen im europäischen Forschungsraum schaffen wird. Open Access bedeutet damit den freien Zugang sowohl zu wissenschaftlichen Publikationen als auch zu Forschungsdaten im Internet. […] Das vorrangige Ziel von OpenAIREplus ist somit die Schaffung einer robusten, kollaborativen Dienstleistung für die Vernetzung von wissenschaftlichen Publikationen, die durch Peer Review-Verfahren begutachtet wurden, und den damit verbundenen Daten. Das Projekt wird eine webbasierte Infrastruktur schaffen, die Informationen von Open Access zugänglichen Forschungsdatensätzen sammelt, anreichert und bereitstellt sowie das Management und die Vernetzung der wissenschaftlichen Daten unterstützt." ***idw Pressemitteilung*** http://idw-online.de/de/news457432
P	
Primärdaten	=Rohdaten, =Forschungsdaten
Projektbericht	Zwischen- oder Abschlussbericht des Projektleiters an den Mittelgeber
R	
Regeln guter wissenschaft-licher Praxis	siehe Positionen
Repositorien (engl. Repository)	s. Archivserver Die DFG verwendet auch den Begriff „Forschungsdatenrepositorien", s.: ***Aufforderung zur Antragstellung Informationsmanagement, Ausschreibung „Informationsinfrastrukturen für Forschungsdaten", 28.04.2010, S.2*** http://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_forschungsdaten_1001.pdf
Rohdaten	=Primärdaten, =Forschungsdaten

Positionen zum Umgang mit Forschungsdaten und Forschungsergebnissen

Allianz der deutschen Wissenschaftsorganisationen

„Mit der **Schwerpunktinitiative "Digitale Information"** haben Alexander von Humboldt-Stiftung, Deutsche Forschungsgemeinschaft (DFG), Deutscher Akademischer Austausch Dienst (DAAD), Fraunhofer-Gesellschaft, Helmholtz-Gemeinschaft, Hochschulrektorenkonferenz (HRK), Leibniz-Gemeinschaft, Max-Planck-Gesellschaft und Wissenschaftsrat im Sommer 2008 eine gemeinsame **Initiative zur Verbesserung der Informationsversorgung** in Forschung und Lehre gestartet. Während der Laufzeit der Initiative von 2008 bis 2012 sollen

- der **weitest mögliche Zugang zu digitalen Publikationen, digitalen Daten und Quellenbeständen** gewährleistet werden,
- optimale Vorraussetzungen für die Verbreitung und Rezeption von Publikationen aus deutscher Forschung mithilfe der digitalen Medien geschaffen werden,
- die **langfristige Verfügbarkeit** der weltweit erworbenen digitalen Medien und Inhalte sowie ihre Integration in die digitale Forschungsumgebung sichergestellt werden,
- netzbasierte Formen des wissenschaftlichen Arbeitens durch innovative Informationstechnologien unterstützt werden.“

s. Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen: http://www.allianzinitiative.de/de/ (letzter Zugriff: 08.08.2012)

Siehe hier die Grundsätze im „**Handlungsfeld Forschungsdaten“**:
http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsaetze/ (letzter Zugriff: 08.08.2012)

„Grundsätze (2010)

Präambel Qualitätsgesicherte Forschungsdaten bilden einen Grundpfeiler wissenschaftlicher Erkenntnis und können unabhängig von ihrem ursprünglichen Erhebungszweck vielfach Grundlage weiterer Forschung sein. Dies gilt namentlich für die Aggregation von Daten aus unterschiedlichen Quellen zur gemeinsamen Nutzung. Die nachhaltige Sicherung und Bereitstellung von Forschungsdaten dient daher nicht nur der Prüfung früherer Ergebnisse, sondern in hohem Maße auch der Erzielung künftiger Ergebnisse. Sie bildet eine strategische Aufgabe, zu der Wissenschaft, Politik und andere Teile der Gesellschaft gemeinsam beitragen müssen. Mit dem Ziel, die Qualität, Produktivität und Wettbewerbsfähigkeit der Wissenschaft zu fördern, verabschiedet die Allianz der deutschen Wissenschaftsorganisationen daher folgende Grundsätze für ein koordiniertes weiteres Vorgehen.

Sicherung und Zugänglichkeit In Übereinstimmung mit wichtigen internationalen Organisationen auf dem Gebiet der Förderung und Durchführung von Forschungsaufgaben([1]) unterstützt die Allianz die **langfristige Sicherung und den grundsätzlich offenen Zugang zu Daten aus öffentlich geförderter Forschung.**

Dabei sind die wissenschaftlichen und rechtlichen Interessen der Forscherinnen und Forscher, der Schutz persönlicher Daten von Probanden, Patienten und anderen von den erhobenen Daten betroffenen Personen sowie weitere Verpflichtungen gegenüber Dritten – etwa Kooperationspartnern - zu beachten. Die Grundsätze der guten wissenschaftlichen Praxis finden Anwendung([2]).

Unterschiede der wissenschaftlichen Disziplinen Formen und Bedingungen des Zugangs zu Forschungsdaten müssen gesondert für die jeweiligen Fachdisziplinen unter Berücksichtigung der Art und Weise der Datenerhebung, des Umfangs und der Vernetzbarkeit des Datenmaterials sowie der praktischen Brauchbarkeit der Daten entwickelt werden. Zugleich ist den jeweiligen Lebenszyklen und Nutzungsszenarien der Daten in dem konkreten Forschungsfeld Rechnung zu tragen.

Wissenschaftliche Anerkennung Die Bereitstellung von Forschungsdaten zur weiteren Nutzung ist eine Leistung, die der Wissenschaft als Ganzer zu Gute kommt. Die Allianz ermutigt zur Anerkennung und Förderung dieses zusätzlichen zeitlichen und finanziellen Aufwands.
Lehre und Qualifizierung Für die Forschenden muss ein angemessenes Ausbildungs- und Unterstützungsangebot zum professionellen Datenmanagement bereitgestellt werden, das an den spezifischen Erfordernissen der Disziplinen ausgerichtet ist.
Verwendung von Standards Eine sachgerechte Nutzung von Forschungsdaten setzt voraus, dass sie in standardisierter Form bereitgestellt, dokumentiert und ausreichend mit Metadaten versehen werden. Standards, Metadatenkataloge und Registries sind unter Berücksichtigung fachspezifischer Anforderungen so zu entwickeln, dass auch eine interdisziplinäre Nutzung möglich ist.
Entwicklung von Infrastrukturen Ein nachhaltiges Forschungsdaten-Management stellt vielfältige technische und organisatorische Anforderungen. Diese Anforderungen müssen in **Zusammenarbeit von Wissenschaftlerinnen und Wissenschaftlern mit Informationsspezialistinnen** und Informationsspezialisten definiert werden. Infrastrukturen sind gemäß diesen Anforderungen zu entwickeln und möglichst von Beginn an in internationale und interdisziplinäre Netzwerke interoperabel einzubinden.“

2011
Empfehlungen der Kommission Zukunft der Informationsinfrastruktur im Auftrag der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder: Gesamtkonzept für die Informationsinfrastruktur in Deutschland:
http://www.allianzinitiative.de/fileadmin/user_upload/KII_Gesamtkonzept.pdf (letzter Zugriff: 08.08.2012)

S.29:
„**(7) Forschungsdaten**
Forschungs-Daten sind als nationales Kulturgut anzusehen und sollten im Sinne einer öffentlichen Aufgabe dauerhaft gesichert und der (Fach-)Öffentlichkeit sowie zukünftigen (Forscher-) Generationen zur Nachnutzung zur Verfügung gestellt werden. Gegenwärtig steigt das Volumen an Forschungsdaten dramatisch an; sie werden durch aufwändige Erhebungsverfahren und meist unter Einsatz erheblicher öffentlicher Mittel gewonnen.
Die Erwartungen der Wissenschaftlerinnen und Wissenschaftler als Nutzer betreffen v. a. den Aufbau und die Verstetigung von disziplingetriebenen, nutzerorientierten Infrastrukturen.
Dabei geht es um vier Schwerpunkte: nachhaltige Sicherung, Erschließung/Bereitstellung, Nachnutzung und langfristige Bewahrung von Forschungsdaten.
Es ist eine nationale Allianz zwischen allen Akteuren nötig: Die Wissenschaftler als Datenproduzenten und Datennutzer, die Hochschulen und Forschungseinrichtungen, die Forschungsförderer, die Datenzentren und Infrastruktureinrichtungen sowie Bund und Länder müssen kooperieren, um gemeinsam den Herausforderungen zu begegnen und Deutschland anschlussfähig für den europäischen und internationalen Wettbewerb zu halten. Dabei sollten alle Aktivitäten, bei denen es disziplinspezifisch oder -übergreifend geboten ist, international vernetzt werden.“

S. 29-30
„(8) Informationskompetenz/Ausbildung
[…]
Die Anforderungen in den anderen Handlungsfeldern bedingen, den Begriff der Informationskompetenz um den Aspekt der Ausbildung und neuer Ausbildungsinhalte zu erweitern. **Somit schließt der Begriff alle Bereiche des Arbeitens mit digitaler Information ein**, und zwar sowohl diziplinspezifisch als auch disziplinunabhängig. Er umfasst nicht nur eine allgemeine Recherchekompetenz, sondern bezieht sich auch auf den Umgang mit der Information bis hin zum

elektronischen Publizieren inkl. der Nutzung der besten Lizenz (z. B. *Creative Commons*), ebenso wie den Aufbau und die Nutzung von Virtuellen Forschungs- und Lernumgebungen **sowie den Umgang mit Forschungsdaten, ihre Beschreibung, Speicherung und Langfristarchivierung**."

s. auch den Abschnitt „Handlungsbedarf und Ressourcen": (7) Forschungsdaten, S. 43-44.

BMBF

s. **Forschung transparent: Recherchen nach Forschern und Vorhaben**:
http://www.bmbf.de/de/2762.php (letzter Zugriff: 08.08.2012)
„Die **Daten der Projektförderung** [=Forschungsprojektdaten; d. Verf.] sind für jedermann zugänglich im Förderkatalog des BMBF dokumentiert. Dort sind die Daten zu allen laufenden und den in den letzten Jahren abgeschlossenen Vorhaben versammelt - insgesamt über 90.000 Projekte.

Von den einzelnen Projekten aus gelangt man mit einem Mausklick zu den **Projektberichten** mit den Ergebnissen der jeweiligen Vorhaben. Alle Projektberichte aus BMBF-Projekten sind in der ca. 7 Mio. Titeldaten umfassenden Datenbank der Technischen Informationsbibliothek TIB Hannover recherchierbar. Seit einiger Zeit werden diese Berichte auch im PDF-Format zur Verfügung gestellt. Soweit digital verfügbar, läßt sich jeder Projektbericht mit einem weiteren Mausklick lesen. Für Wissenschaft und Forschung ist diese Nutzung kostenlos. Ältere Berichte in Papierform lassen sich online über das Bestellsystem https://getinfo.de/app- ggf. gegen Erstattung anfallender Kosten - bestellen."

DFG

1998
Merkblatt Verwendungsrichtlinien für Sonderforschungsbereiche mit Regeln guter wissenschaftlicher Praxis (DFG-Vordruck 5.01 – 01/11):
http://www.dfg.de/formulare/5_01/5_01.pdf (letzter Zugriff: 08.08.2012)

S. 1: „Nach Beschluss der Mitgliederversammlung der DFG vom **17.6.1998** sind bei der Inanspruchnahme von Mitteln der DFG die **Grundsätze zur Sicherung guter wissenschaftlicher Praxis einzuhalten**. Hochschulen und andere Forschungseinrichtungen, die weiterhin DFG-Mittel in Anspruch nehmen möchten, **müssen an ihrer Einrichtung entsprechend der Empfehlungen 1 bis 8 Regeln zur Sicherung guter wissenschaftlicher Praxis etablieren**.
Forschungseinrichtungen, die nicht die Empfehlungen 1 bis 8 implementieren oder sich nicht an die Empfehlungen 1 bis 8 halten, **können ab dem 1. Juli 2002 keine Fördermittel mehr bei der DFG beantragen**." (DFG, .01 – 01/11)

„Empfehlung 1
Regeln guter wissenschaftlicher Praxis müssen - allgemein und nach Bedarf spezifiziert für die einzelnen Disziplinen - Grundsätze, insbesondere für die folgenden Themen, umfassen:

- allgemeine Prinzipien wissenschaftlicher Arbeit, zum Beispiel:
- …
- Resultate zu dokumentieren,
- …
- die Sicherung und Aufbewahrung von **Primärdaten**,
- wissenschaftliche **Veröffentlichungen**."

…

S.12: Empfehlung 7

„Primärdaten als Grundlagen für Veröffentlichungen müssen auf haltbaren und gesicherten Trägern in der Institution, wo sie entstanden, für **zehn Jahre** aufbewahrt werden."
…

2006
DFG Positionspapier: Wissenschaftliche Literaturversorgungs- und Informationssysteme – Schwerpunkte der Förderung bis 2015:
http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier.pdf (letzter Zugriff: 08.08.2012)

s. dort S. 8 Aktionsplan Punkt 14:
„Weiterentwicklung der Strukturen zur Bereitstellung von Forschungsprimärdaten"

2009
DFG - Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme Unterausschuss für Informationsmanagement: „Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten"
http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901.pdf (letzter Zugriff: 08.08.2012)

S.2-3:
„In jedem Fall sollten die erhobenen Daten nach Abschluss der Forschungen **öffentlich zugänglich und frei verfügbar** sein.
[…]
Für die Speicherung und Bereitstellung von **Forschungsprimärdaten** müssen **technische und organisatorische Voraussetzungen** gegeben sein, die **aus den einzelnen Fachdisziplinen heraus zu entwickeln** sind.
[…]
Für die Schaffung der erforderlichen organisatorischen Voraussetzungen und Strukturen sollen **fachspezifische Arbeitsgruppen** eingesetzt werden, die für ihre Community Vorschläge zur Bestimmung der an der Speicherung und Bereitstellung beteiligten Institutionen sowie der Verfahren und Strukturen erarbeiten.
[…]
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.
[…]
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**.
…

7 Empfehlungen
[…]
Die Wissenschaftler sollen in ihren Fachcommunities selber entscheiden, ob bereits Rohdaten hierzu zählen oder ab welchem Grad der Aggregierung die Daten langfristig aufzubewahren sind. Des Weiteren soll die Granularität in groben Umrissen vereinbart sein: wie viele Daten ergeben einen Datensatz, der mit einer stabilen Adresse (persistent identifier) ausgestattet wird?
…
Es ist ein fachspezifisches Organisationskonzept festzulegen, das die nachhaltige Speicherung der Daten regelt.
[…]

Die Speicherung der Forschungsdaten erfolgt im Rahmen definierter Standards
[…]
Die Daten werden durch Metadaten beschrieben. (mindestens nach Dublin Core)“
…

2010
DFG Merkblatt für die Beantragung von Sachbeihilfe (DFG-Vordruck 54.01 – 1/12):
http://www.dfg.de/formulare/54_01/54_01_de.pdf (letzter Zugriff: 08.08.2012)

S.6:
„2.4 Umgang mit den im Projekt erzielten Forschungsdaten
Die Verbesserung des Umgangs mit Forschungsdaten hat sowohl bei nationalen und internationalen Forschungsorganisationen, als auch in der Wissenschaft eine hohe Be-deutung. Die DFG ist daher bestrebt, durch ihre Förderung auch zur Sicherung, Auf-bewahrung und nachhaltigen Verfügbarkeit der Forschungsdaten beizutragen.

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. Bitte berücksichtigen Sie dabei auch - sofern vorhanden - die in Ihrer Fachdisziplin existierenden Standards und die Angebote bestehender Datenrepositorien.“

DFG Antragsmuster für die Einrichtung eines Sonderforschungsbereichs (DFG-Vordruck 60.100.-06/12):
http://www.dfg.de/formulare/60_100/60_100_de_rtf.rtf (letzter Zugriff: 08.08.2012)
S.11:
„1.3.3 Forschungsdaten
<Die Verbesserung des Umgangs mit Forschungsdaten hat sowohl bei nationalen und internationalen Forschungsorganisationen als auch in der Wissenschaft eine hohe Bedeutung. Die DFG ist daher bestrebt, durch ihre Förderung auch zur Sicherung, Aufbewahrung und nachhaltigen Verfügbarkeit der Forschungsdaten beizutragen.

Wenn im Rahmen des geplanten Sonderforschungsbereichs 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 ergriffen werden sollen, um die Daten nachhaltig zu sichern und ggf. für eine erneute Nutzung bereit zu stellen. Bitte berücksichtigen Sie dabei auch – sofern vorhanden – die existierenden Standards und die Angebote bestehender Datenrepositorien. Stellen Sie bitte auch dar, durch welche Einrichtungen (Datenkurator, Rechenzentrum, Bibliothek, etc.) welche Form von Unterstützung beim Daten- und Informationsmanagement durch die am Sonderforschungsbereich beteiligten Institutionen geleistet werden soll.

S.20:
„C 3.4 Planung des Teilprojekts zur Informationsinfrastruktur
Ein Serviceprojekt zur Informationsinfrastruktur kann sich auf eines oder mehrere der im Folgenden erläuterten Ziele beziehen.
Aufbau einer Datenbank zur Speicherung der im Sonderforschungsbereich anfallenden Forschungsdaten einschließlich der Vergabe von Metadaten. …
Pflege und Erschließung von Forschungsdaten. Gefördert werden können die Implementierung und Entwicklung von Techniken und Verfahren zur Pflege und Erschließung von Informationen sowie zur Verknüpfung der Forschungsdaten, die im Sonderforschungsbereich gewonnen werden, mit anderen

Datensystemen auch außerhalb des Verbundes oder eine Einbettung in diese (Referenzierung von Daten).
Nachnutzung von Forschungsdaten. Das Datenspeicherungssystem ist so aufzubauen, dass die Nachnutzbarkeit der im Sonderforschungsbereich erzeugten Forschungsdaten ermöglicht wird. Dies kann etwa durch die Implementierung oder Entwicklung von Interoperabilität bzw. Schnittstellen der Retrieval- und Referenzierungsverfahren oder sonstigen Softwarelösungen und Techniken mit Datenbanksystemen (Fachdatenbanken, Online-Kataloge u.ä.), die außerhalb des Verbundes bestehen und von den jeweils einschlägigen Disziplinen genutzt werden, geschehen.
[…]
Im Teilprojekt ist sicherzustellen, dass neben der Expertise zum Datenmanagement auch datentechnische bzw. informatische Fachkenntnisse eingebunden sind. **Die systematische Zusammenarbeit des Sonderforschungsbereichs mit Informationseinrichtungen (z.B. Bibliotheken, Rechenzentren) muss gesichert sein.** In der Regel soll mit den einschlägig ausgewiesenen Informationseinrichtungen der antragstellenden Hochschule zusammengearbeitet werden. Diese Kooperationen sind zu belegen.“

2012

Positionspapier der Deutschen Forschungsgemeinschaft
Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme
Die digitale Transformation weiter gestalten – Der Beitrag der Deutschen Forschungsgemeinschaft zu einer innovativen Informationsinfrastruktur für die Forschung (3.7.2012)

http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier_digitale_transformation.pdf (letzer Zugriff: 10.09.2012)

S.2:
„Ziel des Systems ist ein für Nutzerinnen und Nutzer freier und umfassender Zugang zu digitaler wissenschaftlicher Information sowie die Vernetzung von Wissen und Daten. Dies setzt zugleich voraus, dass geeignete rechtliche, technische und finanzielle Rahmenbedingungen geschaffen werden, die ein optimales wissenschaftliches Arbeiten ermöglichen.“

S.2 f:
„Erstens wird in jeder Phase der Weiterentwicklung eine strikte Ausrichtung an den Interessen der Wissenschaftlerinnen und Wissenschaftler erfolgen. Das setzt einen intensiven Dialog aller Beteiligten voraus.
Zweitens zielt ein abgestimmtes System von Informationsinfrastrukturen für die Wissenschaft nicht allein auf die Bereitstellung von Information. Es bietet darüber hinaus sowohl die Möglichkeit, digitale Inhalte umfassend nach zu nutzen als auch neue Formen des kooperativen wissenschaftlichen Arbeitens, beispielsweise im Rahmen von Virtuellen Forschungsumgebungen, zu unterstützen. Vor diesem Hintergrund ist es unabdingbar, dass die digitalen Inhalte sowohl bezogen auf ihre technischen als auch ihre rechtlichen Eigenschaften in einem umfassenden Sinne zugänglich und nachnutzbar sind.
Drittens schließlich reicht ein leistungs- und zukunftsfähiges System von Informationsinfrastrukturen über regionale und nationale Grenzen hinaus. Die Anschlussfähigkeit an und Einbindung in internationale Strukturen sind daher immer mit zu berücksichtigen.“
S.4:
„Alle vorgeschlagenen Maßnahmen zielen darauf ab, den möglichst offenen Zugang zu wissenschaftlich relevanter Information, zu Forschungsdaten sowie zu Arbeits- und Kommunikationsplattformen zu optimieren. Dabei setzt sich die DFG konsequent für das Paradigma des Open Access ein.“

S. 17 f.:
„*Forschungsdaten.* Während einige Disziplinen Forschungsdaten bereits seit Längerem teilen und nachnutzen und über entsprechende Strukturen und Praxiserfahrungen verfügen, werden organisatorische Grundlagen und Forschungsdatenrepositorien in anderen Disziplinen derzeit erst entwickelt und aufgebaut. Dabei ist einem breiten Verständnis von Die digitale Transformation weiter gestalten – DFG-Positionspapier Seite 18 von 18
Forschungsdaten auch im Sinne von Informationen zu Objekten und wissenschaftlichen Sammlungen Rechnung zu tragen. Dem einerseits großen Interesse und Bedarf an geeigneten Lösungen stehen gleichzeitig zahlreiche offene Fragen bezogen auf Organisationsformen und Verantwortlichkeiten, rechtliche Rahmenbedingungen und die Sicherstellung von Nachhaltigkeit gegenüber. Hinsichtlich der Forschungsdaten ist bei zukünftigen Fördermaßnahmen besonderes Augenmerk auf die fachspezifisch unterschiedlichen Anforderungen zu richten. Angesichts des hohen Bedarfs gilt es, Fördermaßnahmen in geeigneter Weise zügig zu gestalten und anzubieten.“

DINI - Deutsche Initiative für Netzwerkinformation

2009

Arbeitsgruppe „Elektronisches Publizieren“: Positionspapier Forschungsdaten: http://edoc.hu-berlin.de/series/dini-schriften/2009-10/PDF/10.pdf (letzter Zugriff: 08.08.2012)

S. 5:
„Weltweit gewinnt die Forderung nach einem verantwortungsvollen und organisierten Umgang mit Forschungsdaten, die im Rahmen der öffentlich geförderten Forschung entstehen, an Bedeutung. Wissenschaftliche Einrichtungen sind gefordert sich dieser Herausforderung zu stellen. Besondere Bedeutung kommt dabei der engen Zusammenarbeit zwischen Forschenden und Infrastruktureinrichtungen, wie Bibliotheken, Daten und Rechenzentren zu. Dabei gilt es, kooperativ Lösungen und Wege für das Management von Forschungsdaten zu etablieren, die den Anforderungen der jeweiligen Disziplin entsprechen.“

S.6:
„Um eine Nachnutzung der Forschungsdaten zu ermöglichen, ist es unverzichtbar, organisatorische und technische Strukturen zu schaffen. Während in einigen Disziplinen wie z. B. in den Bio-, Geo-, und Sozialwissenschaften schon seit längerer Zeit Institutionen etabliert sind, die die Aufgaben des Forschungsdatenmanagements übernehmen, fehlen diese in anderen Fachbereichen noch.“

EU

2008

European Commission: Open Access Pilot in FP7: http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-pilot_en.pdf (letzter Zugriff: 08.08.2012)

S.2:
“The European Commission is conducting a pilot initiative on open access to **peer reviewed research articles** in its Seventh Research Framework Programme (FP7). In this pilot, **open access to articles** resulting from research funded in areas participating in the pilot should be provided within a specified time period. This pilot will run until the end of FP7.
Scope of open access pilot
This initiative covers approximately 20% of the FP7 budget and applies to the following areas:
• Energy;

• Environment;
• Health;
• Information and Communication Technologies (Challenge 2: Cognitive Systems, Interaction, Robotics);
• Research Infrastructures (e-Infrastructures);
• Science in Society;
• Socio-economic Sciences and Humanities."

The European Commission's Open Access Pilot for Research Articles: Frequently Asked Questions: http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/08/548&format=HTML&aged=0&language=EN&guiLanguage=en (letzter Zugriff: 08.08.2012)

"Open access to what?
The Commission's open access pilot targets **peer reviewed scientific journal articles** that result from EU funded research.
However, the concept of open access can also apply to **research data, images, etc.**
[…]

How will the open access pilot be implemented?
New grant agreements in the areas covered by the pilot will contain a clause requiring grant recipients to deposit peer reviewed research articles or final manuscripts resulting from their FP7 projects into their institutional or if unavailable a subject-based repository. They will have to make their best efforts to ensure open access to these articles within six or twelve months after publication, depending on the research area.
[…]

European Commission: Open Access in FP7: http://ec.europa.eu/research/science-society/index.cfm?fuseaction=public.topic&id=1300

"This pilot is supported and monitored through the OpenAIRE project. During and after FP7, the Commission will analyse the impact of the pilot and its benefits to EU-funded research with the European Parliament, Member States and stakeholders. The Open Access pilot could potentially be used as a model for the next Framework Programme, and in the Member States."

EU-Projekt: OpenAIRE: „Where to deposit: http://www.openaire.eu/open-access/open-access-in-fp7#? (letzter Zugriff: 08.08.2012)
"Researchers should deposit final articles or manuscripts into the institutional repository of the research institution with which they are affiliated. If this is not possible, they should identify an appropriate subject based/thematic repository. **The Commission provides a special repository for articles that can be stored neither in institutional nor in subject-based/thematic repositories.**
[…]
the Orphan Repository (letzter Zugriff: 08.08.2012) for articles that can be stored neither in institutional nor in subject-based/thematic repositories."

Europäische Kommission, 17.07.2012:
„Verbesserung des Zugangs zu wissenschaftlichen Informationen: …": http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0401:FIN:DE:PDF (letzter Zugriff: 11.09.2012)

S.7:
„Ein breit angelegter, gerechter, tragfähiger und einfacher Zugang zu wissenschaftlichen Informationen, die durch öffentliche Mittel zustande gekommen sind, und ihre nachhaltige

Bewahrung im Hinblick auf eine Weiterverwendung können wesentlich zur Steigerung des wirtschaftlichen Wachstums von Europa beitragen und die Bewältigung der gesellschaftlichen Herausforderungen des 21. Jahrhunderts erleichtern."

EUROHORCs / ESF

EUROHORCs / ESF (2009): The Road Map for Actions to Help Construct the ERA:
http://www.eurohorcs.org/SiteCollectionDocuments/ESF_Road%20Map_long_0907.pdf **(**letzter Zugriff: 08.08.2012)

S.16/17:
„The collection of **research data** is a huge investment. Permanent access to such data, if quality controlled and in interoperable formats, will allow better use to be made of this investment because it allows other researchers to (re)use them. Furthermore it allows re-analysis and could play a role in ensuring research integrity."

S.19:
„Permanent preservation and open access, such as promoted by the Alliance for Permanent Access, will be the rule for repositories. EUROHORCs and ESF Member Organisations will address how to best promote and ensure such permanent access to data generated with their funding."

OECD

2007
OECD: OECD Principles and Guidelines for Access to Research Data from Public Funding:
http://www.oecd.org/dataoecd/9/61/38500813.pdf (letzter Zugriff: 08.08.2012)

S.13:
„These Principles and Guidelines […] provide broad policy recommendations to the governmental science policy and funding bodies of member countries on **access to research data from public funding**. They are intended to promote data access and sharing among researchers, research institutions, and national research agencies, while at the same time, recognising and taking into account, the various national laws, research policies and organisational structures of member countries."

Jens Klump: Digitale Forschungsdaten, in: **nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, hg. v. H. Neuroth, A. Oßwald, R. Scheffel, et al., Göttingen, 2010 [Version 2.3] :** http://nestor.sub.uni-goettingen.de/handbuch/artikel/nestor_handbuch_artikel_427.pdf (letzter Zugriff: 08.08.2012

S.107:
„…im Dezember 2006 verabschiedete der Rat der OECD eine **„Empfehlung betreffend den Zugang zu Forschungsdaten aus öffentlicher Förderung". Diese Empfehlung ist bindend** und muss von den Mitgliedsstaaten der OECD in nationale Gesetzgebung umgesetzt werden, die Umsetzung wird von der OECD beobachtet. In Abschnitt M der Empfehlung wird vorgeschlagen, dass schon bei der Planung von Projekten eine nachhaltige, langfristige Archivierung der Daten berücksichtigt wird."

Wissenschaftsrat

2012

Wissenschaftsrat, 13.07.2012, "Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020":
http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf (letzter Zugriff: 10.09.2012)

S.53 f.:
„Daraus ergeben sich hohe Anforderungen an die Qualität der Daten, ihre Beschreibung und Ergänzung durch standardisierte Metadaten sowie die Sicherung ihrer langfristigen Zugänglichkeit. Der Wissenschaftsrat tritt für maximale Publizität digitaler Forschungsdaten ein und schließt sich den Ausführungen und Empfehlungen der Allianz-Initiative „Digitale Information“ sowie der KII zu den Forschungsdaten an, die auch zum open access Stellung nehmen. Er hält jedoch die folgenden ergänzenden und vertiefenden Hinweise und Empfehlungen für erforderlich.“
[…]
„Vor diesem Hintergrund ist eine **wirksame Verlässlichkeitskontrolle** der Daten und somit auch der mit ihnen generierten Forschungsergebnisse unabdingbar. Dies erfordert insbesondere die **Transparenz der Datenherkunft**, die durch die Metadaten ermöglicht werden muss. Ebenso wie die Nachnutzung von Forschungsdaten setzt die wissenschaftliche Qualitätssicherung die Arbeit von Expertinnen und Experten bei der Bereitstellung, die Vermittlung dieser Anforderung in Studium, Lehre und Nachwuchsförderung sowie die Sicherung und Zugänglichkeit der Daten für Wissenschaftlerinnen und Wissenschaftler voraus.“

S. 55:
„Neben dem fehlenden Zugang stellt vielfach die mangelhafte Dokumentation ein großes Hindernis dar für eine wissenschaftliche Nutzung der teilweise unter erheblichem finanziellem Aufwand gewonnenen Daten sowie vieler born digitals.“
[…] „dass es für viele Forschungsbereiche an geeigneten Repositorien mangelt.“

S. 56:
„Die Fachgemeinschaften bzw. Akteure in interdisziplinären Forschungsfeldern sollten, soweit nicht bereits vorhanden, **Qualitätskriterien** für die Generierung von Forschungsdaten und Richtlinien für ein geeignetes Datenmanagement erarbeiten. Diese sollten insbesondere Konventionen für die **Erstellung von Metadaten** sowie Hinweise zum **Umgang mit datenschutzrechtlich oder ethisch sensiblen Daten** enthalten. Überdies sollten sie die Datenproduzenten sowie die Betreiber von Informationsinfrastrukturen dazu verpflichten**, identifizierte Fehler in den Datensätzen zu korrigieren**. Um die für die wissenschaftliche Qualitätssicherung und die vermehrte Nachnutzung von Daten grundlegende Bereitschaft zu fördern, Forschungsdaten einer wissenschaftlichen Öffentlichkeit zugänglich zu machen, muss die Erhebung und Publikation von Forschungsdaten als eigenständige Forschungsleistung anerkannt werden. Der Wissenschaftsrat spricht sich dafür aus, Forschungsdaten zu veröffentlichen und dabei auch open access-Publikationen zu nutzen. Digital publizierte Forschungsdaten müssen mit so genannten **persistenten Identifikatoren** (z. B. DOI) ausgestattet werden, um sie in wissenschaftlichen Publikationen zitierfähig zu machen.[…]“

S.57:
„Die Forschungsdatenzentren können direkt bei den Datenproduzenten (auch an Hochschulen), bei Archiven oder Forschungseinrichtungen mit Archivfunktionen eingerichtet sein.“