

# Informationsinfrastrukturangebote für digitale Forschungsdaten

Heinz Pampel & Maxi Kindling

*Beitrag zum E(hren)-Journal für Peter Schirmbacher anlässlich seiner Emeritierung als Professor am Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin, herausgegeben von Boris Jacob, Maxi Kindling und Uwe Müller. Das E(hren)-Journal ist online verfügbar unter: <http://ehrenjournal.ib.hu-berlin.de/> Der Text ist online verfügbar unter: <http://nbn-resolving.de/urn:nbn:de:kobv:11-100244024>. Er steht unter der Creative-Commons-Lizenz mit Namensnennung und Weitergabe unter gleichen Bedingungen (CC BY-SA): <http://creativecommons.org/licenses/by-sa/4.0/deed.de>.*

## Zusammenfassung

Mit dem Internet und der Digitalisierung haben sich für die Wissenschaft neue Möglichkeiten der Verbreitung und des Austausches von Wissen und Information eröffnet. Unter dem Stichwort Open Science wird neben Open Access auch die Zugänglichkeit und Nachnutzung von digitalen Forschungsdaten diskutiert. Peter Schirmbacher hat sich in Forschung und Lehre mit der Relevanz, der Gestaltung und der Nutzung von Informationsinfrastrukturangeboten für digitale Forschungsdaten befasst und damit wegweisende Beiträge zur Weiterentwicklung dieser Services für die digital arbeitende Wissenschaft geleistet. Vorliegender Beitrag gibt einen Überblick über den Entwicklungsstand der Informationsinfrastrukturangebote für digitale Forschungsdaten und ordnet sie in die wissenschaftspolitische Diskussionen um Open Science ein.<sup>1</sup>

## Einleitung – Diskussion in der Wissenschaft

Mit dem Internet und der Digitalisierung haben sich für die Wissenschaft neue Möglichkeiten der Verbreitung und des Austausches von Wissen und Information eröffnet. Dieses Potenzial wurde 2003 in der „Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities“ umfassend beschrieben (Max-Planck-Gesellschaft (2003)). Die Erklärung, die 2006 auch von der langjährigen Wirkungsstätte von Peter Schirmbacher, der Humboldt-Universität zu Berlin, un-

<sup>1</sup> Dieser Artikel beinhaltet eine aktualisierte und überarbeitete Fassung des Kapitels „Werkzeuge des Forschungsdatenmanagements“, im „Lehrbuch Forschungsdatenmanagement“ unter [https://handbuch.tib.eu/w/Lehrbuch\\_Forschungsdatenmanagement](https://handbuch.tib.eu/w/Lehrbuch_Forschungsdatenmanagement).

terzeichnet wurde<sup>2</sup>, betont über die Bedeutung des Open Access zu textuellen Publikationen (Suber (2012)) hinaus auch die Öffnung und die Nachnutzung von Forschungsdaten für die digital arbeitende Wissenschaft.

Die Diskussion über die Chancen und Herausforderungen des „data sharings“ hat sich über die letzten Jahre deutlich weiterentwickelt und als Bestandteil von Open Science wissenschaftspolitisch manifestiert (Kroes (2012); European Commission (2016)). Die Empfehlung der ehrwürdigen Fachgesellschaft Royal Society an alle Forschenden – „Where data justify it, scientists should make them available in an appropriate data repository.“ – in der breit rezipierten Denkschrift „Science as an open enterprise“ unterstreicht die Forderung der Wissenschaft nach der offenen Zugänglichkeit von Forschungsdaten. Dieser Wunsch wird durch vielfältige Erhebungen bestätigt (Kuipers & Van der Hoeven (2009); Schäfer et al. (2011); European Commission (2012); Tenopir et al. (2011); Tenopir et al. (2015)). Diese Studien beschrieben jedoch auch die rechtlichen, technischen und kulturellen Barrieren, die einzelne Forschende davon abhalten Forschungsdaten zugänglich zu machen. Diese Barrieren sind häufig auf fehlende Serviceangebote von Informationsinfrastrukturen zurückzuführen (Borgman (2007); Tenopir et al. (2011); van der Graaf & Waaijers (2011)) und machen deutlich, dass sich die Praktiken rund um die Zugänglichmachung und Nachnutzung von Forschungsdaten noch in einem recht jungen Entwicklungsstand befinden.

## Relevanz von Informationsinfrastrukturen

Mit Blick auf digital vorliegende Forschungsdaten bedarf es Informationsinfrastrukturen, die wie von Peter Schirnbacher definiert als „Gesamtheit der technischen, organisatorischen, finanziellen und personellen Ressourcen zur Gewährleistung der Informationsversorgung von Wissenschaft und Forschung“ (Schirnbacher (2015)) beitragen, um so z. B. Forschende beim Management der wachsenden Menge an digitalen Daten zu unterstützen. Informationsinfrastruktureinrichtungen sind insbesondere Bibliotheken, Rechenzentren, Datenzentren, Archive und Museen. Als Teil der Forschungsinfrastruktur verstehen sie sich als Dienstleister für eine definierte Nutzergruppe. Ihre Services werden in diesem Beitrag als Informationsinfrastrukturangebote gefasst. Im Zuge der Digitalisierung sind sie bestrebt die von ihnen gespeicherten Informationsobjekte über das Internet zugänglich zu machen um die Transparenz, Reproduzierbarkeit und Effizienz der Forschung zu verbessern.

In der Informationswissenschaft sind die Informationsinfrastrukturen und entsprechende Angebote Gegenstände von Forschung und Entwicklung. So wurden am Lehrstuhl Informationsmanagement von Peter Schirnbacher am Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität in Kooperation mit der Arbeitsgruppe Elektronisches Publizieren der

---

<sup>2</sup> Siehe: „Berliner Erklärung, Signatoren“, zuletzt geprüft am 31.01.2017, <https://openaccess.mpg.de/3883/Signatories>.

Universitätsbibliothek und des Computer- und Medienservice der Humboldt-Universität Anforderungen aus der Praxis erhoben, Angebote wie Repositorien und Mehrwertdienste aufgebaut, analysiert und weiterentwickelt, die technische und organisatorische Vernetzung dieser Infrastrukturen vorangetrieben sowie Qualitätsstandards entwickelt und verbreitet. Darüber hinaus wurden Geschäfts- und Finanzierungsmodelle eruiert und die Anwendung von geeigneten rechtlichen Lizenzmodellen betrachtet (Schirnbacher (2015)).

Die Arbeiten, die Peter Schirnbacher in diesem Bereich geleistet hat, haben in den letzten Jahren einen wichtigen Beitrag für die Weiterentwicklung und Professionalisierung dieser Informationsinfrastrukturangebote geleistet, die das Forschungsdatenmanagement als „gesamte[n] Prozess, der die Allokation, die Generierung, die Bearbeitung und Anreicherung, die Archivierung und Veröffentlichung von digitalen Forschungsdaten selbst oder von einer entsprechenden klassischen Textproduktion unterstützt“ (Kindling & Schirnbacher (2013)). Für die Auseinandersetzung mit der Thematik schlägt Peter Schirnbacher darüber hinaus eine Betrachtungsweise des Forschungsdatenmanagements vor, die zwischen wissenschaftspolitischer, organisatorischer und technischer Dimension unterscheidet.<sup>3</sup>

In Deutschland ist spätestens seit der Initiierung der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen<sup>4</sup> im Jahre 2008 eine steigende wissenschaftspolitische Relevanz dieser Infrastrukturen zu beobachten. Wichtige Arbeiten dazu hat die Enquete-Kommission „Internet und digitale Gesellschaft“ (2012) des 17. Deutschen Bundestags geleistet. Deren Projektgruppe „Bildung und Forschung“ hat sich umfassend mit den Auswirkungen der Digitalisierung in Forschung und Lehre befasst und sich im parteiübergreifenden Konsens dafür ausgesprochen, „das Open Access-Prinzip, soweit möglich, auf den Bereich der Forschungsdaten aus öffentlich geförderter Forschung auszuweiten und hierfür entsprechende Maßnahmen zu erarbeiten“.

Ausgehend von den Grundsätzen zum Umgang mit Forschungsdaten, die die Allianz der Wissenschaftsorganisationen 2010 veröffentlicht hat, haben Bund und Länder auf die Impulse aus der Wissenschaft reagiert und Empfehlungen zur zukünftigen Ausgestaltung dieser Infrastrukturen in Auftrag gegeben, die auf nationaler Ebene durch die Kommission Zukunft der Informationsinfrastruktur (KII) und den Wissenschaftsrat (WR) bearbeitet wurden und im Jahr 2013 zur Einsetzung des Rates für Informationsinfrastrukturen (RfII) durch die Gemeinsame Wissenschaftskonferenz (GWK) geführt haben. Anliegen der Politik ist es, die „vielfältigen Aktivitäten in diesem Bereich besser aufeinander abzustimmen und sie miteinander zu verzahnen“ (Gemeinsamen Wissenschaftskonferenz (2013)). Der RfII legte 2016 die Empfehlung „Leistung aus Viel-

<sup>3</sup> Die Dimensionen wurden im Rahmen der Vorlesung „Ausgewählte Aspekte digitaler Informationsversorgung“ am Institut für Bibliotheks- und Informationswissenschaft erarbeitet. Eine Publikation von Peter Schirnbacher dazu ist derzeit in Vorbereitung.

<sup>4</sup> Siehe: „Allianz der deutschen Wissenschaftsorganisationen“, zuletzt geprüft am 31.01.2017, <http://www.allianzinitiative.de>.

falt“ vor, die im Kern die Schaffung einer „Nationalen Forschungsdaten-Infrastruktur (NFDI)“ beinhaltet (Rat für Informationsinfrastrukturen (2016)).

## Grundlagen der Zugänglichmachung und Nachnutzung von Forschungsdaten

Digitale Forschungsdaten können unter verschiedenen Zugangsbedingungen und -verfahren bereitgestellt werden. In einem ersten Schritt ist es dabei hilfreich, sich auf die Zugangsbereiche von Treloar & Harboe-Ree (2008) zu beziehen, die zwischen einer privaten, einer Gruppen- und einer öffentlichen Domäne unterscheiden. Bei dieser Darstellung handelt es sich um eine konzeptuelle Trennung von Zugangsdomänen, die in der Praxis entsprechende Implikationen z. B. für die Aufbereitung der Forschungsdaten hat. Werden Forschungsdaten in der öffentlichen Domäne geteilt, so sind sie idealerweise durch einen persistenten Identifikator zitierfähig, mit aussagekräftigen Metadaten zu ihrer Entstehung, Verarbeitung und den Nutzungsmöglichkeiten angereichert.

Die Bereitstellung von Forschungsdaten nach den Prinzipien von Open Access, die in der „Berlin Declaration“ aus dem Jahr 2003 niedergeschrieben ist, bedeutet, dass Forschungsdaten ohne technische, rechtliche und finanzielle Einschränkungen zugänglich gemacht werden und nachnutzbar sind (Max-Planck-Gesellschaft (2003)).

Seit einigen Jahren gibt es in der Wissenschaft unter dem Begriff Open Research Data eine Diskussion über die Strategien und Verfahren der Öffnung von Forschungsdaten, die als Teil der Entwicklung hin zu Open Science verstanden wird (Pampel & Dallmeier-Tiessen (2014)).

Um Forschenden Anreize zur Zugänglichmachung ihrer Daten zu schaffen, haben sich im Zusammenspiel von Wissenschaft, Bibliotheken und Verlagen in den vergangenen Jahren Publikationsstrategien etabliert, die den Forschenden, die Forschungsdaten Dritten bereitstellen, entsprechende Anerkennung im Reputationssystem sichern. Zentrale Voraussetzung dafür ist, wie oben erwähnt, ihre Zitierbarkeit durch die dauerhafte Adressierung der Datensätze (Klump et al. (2006); Neumann & Brase (2014)) z. B. durch die Vergabe des Digital Object Identifier (DOI). Rund um den DOI entsteht durch die Möglichkeiten der Verlinkung mit weiteren persistenten Identifikatoren für Personen (Pampel & Fenner (2016)), Organisationen (Cruse et al. (2016)) sowie weiteren Produkten der wissenschaftlichen Arbeit wie z. B. Proben (Lehnert et al. (2011); Conze et al. (2017) ) die Grundlage für ein offenes Netzwerk, das Informationsobjekte sichtbar und nutzbar macht.

Können Forschungsdaten z. B. aus Gründen des Datenschutzes – wie etwa in der Medizin oder den Sozialwissenschaften – nicht offen zugänglich gemacht werden, ist es empfehlenswert, zumindest einen Teil ihrer Metadaten öffentlich sichtbar zu machen, so dass die Forschungstransparenz

gewährleistet wird und die Chance besteht, dass Dritte sie zuordnen und ihren Kontext kennen können. Die Bedingungen der Zugänglichkeit und Nutzung von Forschungsdaten müssen dann mit den jeweiligen Rechteinhaberinnen und -inhabern geklärt werden. Ein weiteres Szenario ist, dass Forschungsdaten erst nach einer Embargofrist zugänglich gemacht werden. So können Forschende ihre Daten bereits auf einem Forschungsdatenrepositorium speichern und mit einem persistenten Identifikator versehen, der dann in einer Textpublikation (z. B. einem Artikel in einem Fachjournal) zitiert werden kann. Die Forschenden haben somit Zeit, die Forschungsdaten umfassend für die eigene Forschungsarbeit auszuwerten, bevor sie dann für die Fachcommunity zugänglich gemacht werden.

Mit den FAIR-Prinzipien – Findable, Accessible, Interoperable und Reusable – (Wilkinson et al. (2016)) wurden im Jahr 2016 Standards geschaffen, die den unterschiedlichen Anforderungen der Disziplinen entsprechen und als Grundlage für den zeitgemäßen Umgang mit Forschungsdaten gelten können. Sie tragen damit auch der Forderung nach einer möglichst offenen Zugänglichmachung der Daten Rechnung. Die FAIR-Prinzipien bilden einen Rahmen für das Forschungsdatenmanagement an wissenschaftlichen Einrichtungen und wurden z. B. von der Europäischen Kommission in ihrer Rolle als Förderorganisation im Forschungsrahmenprogramm Horizon 2020 verankert (European Commission (2016)).

## Leit- und Richtlinien

Unter dem Begriff „Data Policies“ werden Leit- und Richtlinien zum Umgang mit wissenschaftlichen Daten gefasst. Je nach Akteur und Zielgruppe variieren diese Leit- und Richtlinien in ihrer Verbindlichkeit (Pampel & Bertelmann (2011)). Zentral sind in Deutschland die – vor dem Hintergrund eines spektakulären Falles von wissenschaftlichem Fehlverhalten formulierten – „Vorschläge zur Sicherung guter wissenschaftlicher Praxis“ der Deutschen Forschungsgemeinschaft (DFG) aus dem Jahr 1998 (2013 aktualisiert), die an vielen wissenschaftlichen Institutionen in Deutschland bindenden Charakter haben. Sie empfehlen die Aufbewahrung von Forschungsdaten für mindestens zehn Jahre. Die von der DFG geforderte „Sicherung und Aufbewahrung“ stellt jedoch weder die Zugänglichkeit der Forschungsdaten noch deren Nachnutzbarkeit sicher (Deutsche Forschungsgemeinschaft (2013)). Die „Sicherung und Aufbewahrung“ sind somit als Mindestanforderung für den Umgang mit Forschungsdaten in der Wissenschaftspraxis zu verstehen, die zu erweitern ist. Im Jahr 2015 hat die DFG zusätzlich „Leitlinien zum Umgang mit Forschungsdaten“ veröffentlicht, in denen Mittelempfängerinnen und -empfänger aufgefordert werden „Forschungsdaten so zeitnah wie möglich verfügbar“ zu machen und sicherzustellen, so dass „eine sinnvolle Nach- und Weiternutzung durch Dritte ermöglicht“ wird (Deutsche Forschungsgemeinschaft (2015)).

In den letzten Jahren haben viele wissenschaftliche Einrichtungen solche Leit- und Richtlinien verabschiedet.<sup>5</sup> Auch an der Humboldt-Universität wurde auf Basis der Arbeiten am Lehr- und Forschungsbereichs Informationsmanagement von Peter Schirmbacher eine solche „Forschungsdaten-Policy“ auf den Weg gebracht (Humboldt-Universität zu Berlin (2014)). Über diese Leit- und Richtlinien von Förderorganisationen und wissenschaftlichen Einrichtungen hinaus wirken die Vorgaben wissenschaftlicher Zeitschriften auf Forschende. So erwarten z. B. die Zeitschriften von PLOS (Bloom et al. (2014)) und Nature (Nature (2016)) ein sogenanntes „Data Availability Statement“, das bei einer Einreichung über die Zugänglichkeit der dem Artikel zugrundeliegenden Forschungsdaten informiert (Bloom et al. (2014); Nature (2016)).

## Forschungsdatenmanagementpläne

In Forschungsdatenmanagementplänen (im Englischen „Data Management Plan“ genannt) wird der Umgang mit Forschungsdaten in einem Forschungsprojekt von der Planung bis zum Abschluss eines Projekts dokumentiert. Forschungsdatenmanagementpläne sind dynamische Dokumente, die im Projektverlauf immer wieder angepasst werden können. Sie geben einem Projekt und seinen Partnerinnen und Partnern einen Handlungsrahmen für den Umgang mit den im Projekt entstehenden und verwendeten Forschungsdaten. Inhalte von Datenmanagementplänen sind z. B. Aussagen zu Typ und Format der Forschungsdaten, deren Grad der Standardisierung (z. B. verwendete Metadaten, Datenformate), Angaben zur Zugänglichkeit und den Möglichkeiten der Nachnutzung sowie zu den angewandten Verfahren der langfristigen Bewahrung der Forschungsdaten.

Forschungsförderorganisationen fordern zunehmend bereits im Rahmen der Projektantragstellung Aussagen zur Zugänglichkeit und Nachnutzbarkeit der Forschungsdaten. Die konkreten Anforderungen an diese Forschungsdatenmanagementpläne sind in den Leit- oder Richtlinien zur Antragstellung formuliert.

Die Anforderungen der Deutschen Forschungsgemeinschaft (DFG) an Antragstellerinnen und Antragsteller finden sich im „Leitfaden für die Antragstellung“.<sup>6</sup> Dort heißt es: „Wenn aus Projektmitteln systematisch Forschungsdaten oder Informationen gewonnen werden, die für die Nachnutzung durch andere Wissenschaftlerinnen und Wissenschaftler geeignet sind, legen Sie bitte dar, ob und auf welche Weise diese für andere zur Verfügung gestellt werden. Bitte berücksichtigen Sie dabei auch – sofern vorhanden – die in Ihrer Fachdisziplin existierenden Standards und die Angebote existierender Datenrepositorien oder Archive.“ (Deutsche Forschungsgemeinschaft (2014))

<sup>5</sup> Siehe die Übersicht im „Wiki Forschungsdaten.org“, zuletzt geprüft am 31.01.2017, <http://www.forschungsdaten.org/index.php/Data.Policies>.

<sup>6</sup> „DFG, Leitfaden für die Antragsstellung“, zuletzt geprüft am 31.01.2017, <http://www.dfg.de/formulare/54.01/>.

Während die DFG diese Aussagen zum Datenmanagement im Rahmen der Antragstellung erwartet, sind Datenmanagementpläne im Rahmen des europäischen Forschungsrahmenprogramm Horizon 2020 als „deliverable“ in den ersten sechs Monaten eines Projekts einzureichen. Weitere Aussagen eines Projektkonsortiums zum Datenmanagement werden im Projektverlauf und zum Ende eines Projekts erwartet (European Commission (2016)).

Für die Erstellung dieser Pläne gibt es mittlerweile mehrere Werkzeuge und Vorlagen. Beispiele sind:

- Checkliste zum Forschungsdaten-Management des WissGrid-Projekts, Deutschland (Ludwig & Enke (2013)),
- DMPonline des Digital Curation Centre, Großbritannien<sup>7</sup>,
- DMPTool der University of California, U.S.A.<sup>8</sup>

Forschungsdatenmanagementpläne dienen auch als Berechnungsgrundlage für die Kosten des Forschungsdatenmanagements. Teilweise stellen Förderorganisationen Mittel für diese Arbeiten zur Verfügung. Im Rahmen eines Forschungsdatenmanagementplans können die beantragten Mittel so nach Aktivität des Forschungsdatenmanagements aufgeschlüsselt werden.

## Informationsinfrastrukturangebote

Im Folgenden sollen einige der zentralen Angebote der Informationsinfrastruktur beschrieben werden, die das Forschungsdatenmanagement unterstützen und die Forschungstransparenz sowie die Nachnutzung von Forschungsdaten ermöglichen. Sie unterstützen den Nachweis der Forschungsdaten durch beschreibende Metadaten, den Zugang zu den Forschungsdaten für die Öffentlichkeit oder eine ausgewählte Nutzergruppe (z. B. aus datenschutzrechtlichen Gründen) sowie die längerfristige Erhaltung der Forschungsdaten.

In einigen Wissenschaftsdomänen sind Informationsinfrastrukturangebote für Forschungsdaten bereits seit vielen Jahrzehnten etabliert (so z. B. durch die Weltdatenzentren in einigen Disziplinen der Naturwissenschaften schon seit den 1950er Jahren<sup>9</sup>, in anderen wird gerade erst begonnen, entsprechende Angebote aufzubauen und zu etablieren.

Je nach Wissenschaftsdisziplin und davon abhängig nach Methode der Datenerhebung und -verarbeitung sowie nach Datentypen und -formaten bestehen unterschiedliche Anforderungen an die Funktionalitäten der Infrastrukturangebote. Die Landschaft der Infrastrukturangebote

<sup>7</sup> Siehe: „DMPonline“, zuletzt geprüft am 31.01.2017, <http://dmponline.dcc.ac.uk>.

<sup>8</sup> Siehe: „DMPTool“, zuletzt geprüft am 31.01.2017, <http://dmp.cdlib.org>.

<sup>9</sup> Seit 2008 als ICSU World Data System (WDS) fortgeführt. Siehe: „WDS“, zuletzt geprüft am 31.01.2017, <http://www.icsu-wds.org>.

für Forschungsdaten ist daher äußerst heterogen (Pampel et al. (2013)). Seit einigen Jahren wird die Professionalisierung dieser Infrastrukturangebote für die Speicherung, Archivierung und Zugänglichmachung unter dem Begriff „Forschungsdatenrepositorien“ diskutiert. Aus der Perspektive der Wissenschaftspraxis variieren die Publikationsstrategien für Forschungsdaten. Teilweise werden die Daten als eigenständiges Informationsobjekt z.B. in einem Forschungsdatenrepositorium (im Folgenden: FDR) veröffentlicht, teils als „Supplement“ oder „Enhanced Publication“ in direkter Verbindung mit einem Aufsatz, oder auch als „Data Paper“ in einer wissenschaftlichen Zeitschrift oder einem speziellen „Data Journal“ (Woutersen-Windhouwer & Brandsma (2009); (Pampel & Dallmeier-Tiessen (2014)). Ein Konzept für das Zusammenspiel von FDR und Zeitschriftenartikeln wurde 2017 mit dem Scholix Framework vorgelegt (Burton et al. (2017)). Darüber hinaus liegen Empfehlungen zur persistenten Adressierung und Vernetzung der Informationsobjekte für Forschungsdatenrepositorien (Fenner et al. (2016)) und Zeitschriften (Cousijn et al. (2017)) vor, die einen wichtigen Beitrag dazu leisten, dass Forschungsdaten als eigenständige Objekte im Bereich der Wissenschaftskommunikation anerkannt werden.

## Forschungsdatenrepositorien

Der Begriff des „Repositoriums“ ist im Kontext der Informationsinfrastruktur bereits aus dem Umfeld der Textrepositorien bekannt, die Zugang zu elektronischen Veröffentlichungen ermöglichen (Lynch (2003)).

Im Projekt re3data – Registry of Research Data Repositories, an dem der Lehr- und Forschungsbereich von Peter Schirmbacher als Projektpartner beteiligt war, wurde ein globales Verzeichnis von Forschungsdatenrepositorien aufgebaut (Pampel et al. (2013)). In diesem Verzeichnis werden inzwischen weltweit 1.805 FDR nachgewiesen.<sup>10</sup> Sie sind mithilfe eines umfangreichen Metadatenschemas beschrieben (Rücknagel et al. (2015)). In re3data können Forschende, Angehörige von Infrastruktureinrichtungen und Forschungsförderorganisationen sowie Interessierte nach geeigneten FDR für die Zugänglichmachung von Forschungsdaten suchen; zugleich können FDR gefunden werden, die Forschungsdaten zur Nachnutzung anbieten. Auf Basis der umfangreichen Metadatenbeschreibung der indexierten FDR lässt sich die Suche in re3data mithilfe von Filtern verfeinern.

Anhand von re3data zeigt sich die gesamte Bandbreite an Forschungsdatenrepositorien: Es werden beispielsweise disziplinäre, multidisziplinäre, behördliche, institutionelle und projektspezifische Repositorien nachgewiesen.

Für die Aufnahme eines FDR in das Verzeichnis gilt die folgende Definition: „A research data repository is a subtype of a sustainable information infrastructure which provides long-term storage

---

<sup>10</sup> Stand: 26.01.2017



and access to research data. Research data means information objects generated by scholarly projects for example through experiments, measurements, surveys or interviews.“ (Rücknagel et al. (2015)). Diese sehr breit angelegte Definition eines FDR entstand vor dem Hintergrund, dass den Nutzerinnen und Nutzern von re3data bei der Suche möglichst viele potentiell relevante FDR angeboten werden sollen.

Nach informationswissenschaftlichen Gesichtspunkten ist ein FDR spezifischer zu beschreiben: Es handelt sich dabei um ein webbasiertes Angebot der Informationsinfrastruktur, das die Zugänglichkeit und Nutzung von digitalen Forschungsdaten unterstützt. Dabei stellt das FDR mindestens sicher, dass auf Basis von Anforderungen einer definierten Nutzergruppe

- die Forschungsdaten in einem für die Nachnutzung geeigneten Datenformat bereitgestellt werden,
- die Forschungsdaten zitierbar sind und durch etablierte Metadaten schemata beschrieben werden,
- die Forschungsdaten mit Informationen zu Nutzungsbedingungen angereichert werden.

Ein FDR wird durch eine vertrauenswürdige Institution wie etwa eine wissenschaftliche Einrichtung mit dem Ziel betrieben, das FDR langfristig verfügbar zu halten. Die Zielgruppe des FDR, die Anforderungen an die zu speichernden Forschungsdaten, die Bedingungen zur Nutzung des FDR und der Forschungsdaten sind in öffentlich verfügbaren Dokumenten dargestellt. Ein FDR orientiert sich bei der Aufbereitung von Forschungsdaten an in der jeweiligen Fachcommunity verfügbaren Standards sowie an übergreifenden Qualitätsstandards für Informationsinfrastrukturangebote wie etwa dem Data Seal of Approval<sup>11</sup>, der DIN-Norm 31644<sup>12</sup> oder der ISO-Norm 16363:2012 „Audit and certification of trustworthy digital repositories“<sup>13</sup>. Ein FDR bietet eine metadatenbasierte Suche über die gespeicherten Inhalte und erlaubt den Zugang zu diesen über eine Weboberfläche. Darüber hinaus stellt das FDR die nachgewiesenen Metadaten über standardisierte Schnittstellen bereit und erlaubt so Aggregation der Metadaten durch Dienste Dritter (z. B. durch Suchmaschinen und Harvesting-Dienste). Die Gestaltung von FDR ist darüber hinaus abhängig von Faktoren wie z. B.

- den Zugangsmodellen zum Repository selbst und zu den Forschungsdaten,
- den durch jeweilige Zielgruppe benötigten Funktionalitäten,
- der verwendeten Software und
- den zugrundeliegenden Kosten- und Geschäftsmodellen.

<sup>11</sup> Siehe: „Data Seal of Approval“, zuletzt geprüft am 31.01.2017, <http://www.datasealofapproval.org>.

<sup>12</sup> Siehe: „Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644“, zuletzt geprüft am 31.01.2017, <https://www.beuth.de/de/publikation/vertrauenswuerdige-digitale-langzeitarchivierung/169654635>.

<sup>13</sup> Siehe: „ISO 16363:2012“, zuletzt geprüft am 31.01.2017, [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=56510](http://www.iso.org/iso/catalogue_detail.htm?csnumber=56510).

Bekannte fachliche FDR sind PANGAEA (Diepenbroek et al. (2002)) im Bereich der Geowissenschaften und GenBank (Benson et al. (2015)) in der Biomedizin. An vielen wissenschaftlichen Institutionen entstehen aktuell FDR, die Forschungsdaten der Angehörigen der jeweiligen Institutionen aufnehmen. Beispiele in Deutschland sind heiDATA<sup>14</sup> an der Universität Heidelberg, Open Data LMU<sup>15</sup> an der Ludwig-Maximilians-Universität München oder panMetaDocs (Ulbricht et al. (2016)) am Deutschen GeoForschungsZentrum – GFZ. An einigen Einrichtungen wie z. B. der Technischen Universität Berlin<sup>16</sup> wird ein gemeinsames Repositorien für Forschungsdaten, Textpublikationen und andere Informationsobjekte angeboten. Auch an der Humboldt-Universität sind unter der Leitung und Mitwirkung von Peter Schirmbacher in seinen Funktionen als Direktor des Computer- und Medienservice sowie seines Lehr- und Forschungsbereiches am IBI verschiedene Angebote entstanden (wie bspw. das FDR LAUDATIO für historische Textkorpora<sup>17</sup> oder das Medien-Repositorium<sup>18</sup>) bzw. initiiert (z. B. die Nutzung des Publikationsservers edoc für digitale Forschungsdaten im Kontext des von ihm mit beantragten Projekts eDissPlus)<sup>19</sup>.

Ist in der Fachcommunity kein geeignetes FDR vorhanden und auch kein institutionelles FDR existent, bietet sich die Zugänglichmachung der Daten auf einem generischen Repositorium wie beispielsweise auf Zenodo<sup>20</sup> am CERN an.

## Data Papers und Data Journals

In Data Papers werden Forschungsdaten umfassend dokumentiert. Sie eignen sich besonders, wenn die Metadatenbeschreibung eines Datensatzes in einem FDR nicht ausreichend ist, um beispielsweise die Möglichkeiten der Nachnutzung zu dokumentieren. Die Veröffentlichung eines Data Papers erfolgt in einer wissenschaftlichen Zeitschrift, einem speziellen Data Journal, aber auch anderen Publikationstypen (Pampel & Dallmeier-Tiessen (2014))<sup>21</sup>. In den Geowissenschaften wurde 2008 das Data Journal Earth System Science Data (ESSD) gegründet (Pfeiffenberger & Carlson (2011)). Die hohen Zitationsraten der dort dokumentierten Forschungsdaten führten dazu, dass das ESSD in der 2016 Edition des Journal Citation Report (JCR) auf dem Platz drei der meist zitierten geowissenschaftlichen Journale gelistet wurde<sup>22</sup>.

<sup>14</sup> Siehe: „heiDATA“, zuletzt geprüft am 31.01.2017, <https://heidata.uni-heidelberg.de/>.

<sup>15</sup> Siehe: „Open Data LMU“, zuletzt geprüft am 31.01.2017, <http://data.ub.uni-muenchen.de/>.

<sup>16</sup> Siehe: „Deposit Once“, zuletzt geprüft am 31.01.2017, <https://depositonce.tu-berlin.de>.

<sup>17</sup> Siehe: „LAUDATIO Repository“, zuletzt geprüft am 06.02.2017, <http://www.laudatio-repository.org/repository/>

<sup>18</sup> Siehe: „Medien-Repositorium“, zuletzt geprüft am 06.02.2017, <https://medien.hu-berlin.de/>

<sup>19</sup> Siehe: „eDissPlus“, zuletzt geprüft am 06.02.2017, <https://www2.hu-berlin.de/edissplus/>

<sup>20</sup> Siehe: „Zenodo“, zuletzt geprüft am 31.01.2017, <http://zenodo.org>.

<sup>21</sup> Eine Liste von Data Journals wird im Forschungsdaten-Wiki gepflegt: „Data Journals,“ zuletzt geprüft am 31.01.2017, [http://www.forschungsdaten.org/index.php/Data\\_Journals](http://www.forschungsdaten.org/index.php/Data_Journals).

<sup>22</sup> Siehe: „First Journal Impact Factor for ESSD released,“ zuletzt geprüft am 31.01.2017, [http://www.earth-system-science-data.net/about/news\\_and\\_press/2016-06-17\\_first-impact-factor-for-essd.html](http://www.earth-system-science-data.net/about/news_and_press/2016-06-17_first-impact-factor-for-essd.html).

## Recherchemöglichkeiten

Werden Forschungsdaten für die Nutzung gesucht, ist es naheliegend, bekannte und etablierte Angebote wie FDR oder Data Journals direkt anzusteuern oder im jeweiligen Wissenschaftsgebiet z. B. mithilfe von re3data nach FDR zu suchen, die potentiell relevante Forschungsdaten nachweisen. Neben lokalen und fachspezifischen Recherchemöglichkeiten gibt es darüber hinaus auch generische Suchdienste, über die Forschungsdaten recherchiert werden können. Mit Blick auf die zumeist fachspezifischen Metadaten bieten diese Dienste jedoch häufig keine detaillierten Suchmöglichkeiten. Nachfolgend werden drei Beispiele für Dienste genannt, die Metadaten von FDR und anderen Forschungsinfrastrukturen aggregieren und für die angebotsübergreifende Suche anbieten (ihre Datenräume überschneiden sich teilweise). Bei allen Angeboten ist zu beachten, dass die Granularität der über Identifikatoren wie einen DOI identifizierbaren Metadateninformationen sehr unterschiedlich sein kann. Das bedeutet, dass sowohl einzelne Forschungsdaten im Sinne von Dateien als auch ganze Forschungsdatensammlungen mit Metadaten beschrieben sind.

### DataCite Metadata Search<sup>23</sup>

Dieser Dienst erlaubt die Suche in den Metadaten von Informationsobjekten wie z. B. Forschungsdaten, die bei DataCite mit einem Digital Object Identifier registriert sind und für die bei DataCite entsprechende Metadaten hinterlegt sind. Im Februar 2017 sind im Datenraum von DataCite rund 7,6 Mio. Metadatensätze recherchierbar, darunter 3.178.328 unter dem Objekttyp „Dataset“<sup>24</sup>. Diese Angaben sollten nur als Orientierungswerte betrachtet werden, da verschiedene Aspekte zu einer Verfälschung führen könnten (Metadatenqualität, Dubletten, Zuordnung von Nicht-Forschungsdaten als Forschungsdaten).

### BASE – Bielefeld Academic Search Engine<sup>25</sup>

Dieser Dienst erlaubt die Suche in den Metadaten von derzeit 5.135 Quellen wie Repositorien und anderen Datenanbietern der Informationsinfrastruktur, die ihre Metadaten über eine OAI-PMH-Schnittstelle zu Verfügung stellen. Im Februar 2017 sind 3.622.458 Metadateneinträge der Dokumentart „Forschungsdaten“ zugeordnet<sup>26</sup>. Forschungsdaten können darüber hinaus auch durch den jeweiligen Anbieter wie ein FDR auch unter weiteren Dokumentarten kategorisiert sein. In die BASE-Suche sind auch die Metadaten von DataCite integriert.

### EUDAT B2 Find

Dieser Dienst wurde im Rahmen des EU-Projektes EUDAT<sup>27</sup> entwickelt und ermöglicht die Suche in Forschungsdatensammlungen aus diversen ausgewählten Quellen, darunter die EUDAT

<sup>23</sup> Siehe: „DataCite Metadata Search,“ zuletzt geprüft am 31.01.2017, <http://search.datacite.org/ui>.

<sup>24</sup> Stand 06.02.2017

<sup>25</sup> Siehe: „BASE,“ zuletzt geprüft am 31.01.2017, <https://www.base-search.net>.

<sup>26</sup> Stand: 06.02.2017

<sup>27</sup> Siehe: „EUDAT B2 Find,“ zuletzt geprüft am 31.01.2017, <http://b2find.eudat.eu>.

Daten-Zentren<sup>28</sup>. EUDAT B2 Find weist im Februar 2017 472.082 Metadaten von Forschungsdaten nach<sup>29</sup>. Indexiert sind u. a. Metadaten aus „Common Language Resources and Technology Infrastructure – CLARIN“<sup>30</sup> oder „Global Biodiversity Information Facility – GBIF“<sup>31</sup>.

## Ausblick

Unter Mitwirkung von Peter Schirnbacher hat die Hochschulrektorenkonferenz (2014) das „Management von Forschungsdaten als strategische Aufgabe der Hochschulleitungen“ benannt. (Hochschulrektorenkonferenz (2014)) Hochschulen und außeruniversitäre Forschungseinrichtungen haben in den letzten Jahren damit begonnen, Leit- und Richtlinien sowie Informationsinfrastrukturangebote für Forschende rund um die Zugänglichkeit und Nachnutzung von Forschungsdaten zu entwickeln.

Die Schaffung der vom RfII geforderten „Nationale Forschungsdaten-Infrastruktur (NFDI)“ stellt eine große Chance für das deutsche Wissenschaftssystem dar, um die Möglichkeiten der Digitalisierung für Forschung und Lehre zu nutzen (Rat für Informationsinfrastrukturen (2016)) und die sehr differenzierte und heterogene Landschaft der Angebote zusammenzuführen und durch eine gemeinsame Basis zu stützen. Bereits jetzt existierenden über 270 Forschungsdatenrepositorien, an deren Betrieb deutsche Einrichtungen beteiligt sind (Pampel et al. (2016)). Mit Blick auf die Internationalität der Wissenschaft gilt es jedoch den Blick über die Ländergrenzen hinweg zu heben.

Mit dem DFG-Projekt GerDI und der Helmholtz Data Federation (HDF) haben sich erste Akteure ins Spiel gebracht und ihre Bereitschaft zur aktiven Gestaltung der NFDI signalisiert (Helmholtz-Gemeinschaft (2016)); Zentralbibliothek für Wirtschaftswissenschaften (2016)). Beide Akteure betonen die Bedeutung der Anschlussfähigkeit zur European Open Science Cloud (EOSC), deren Schaffung von der Europäischen Kommission vorangetrieben wird (High Level Expert Group on the European Open Science Cloud (2016)). Ein erstes Förderprogramm des Bundesministeriums für Bildung und Forschung (BMBF) auf Bundesebene<sup>32</sup> und diverse Förderaktivitäten auf Länderebene zeigen zwar erste Ansätze zur Realisierung einer NFDI – eine gemeinsame Koordination ist jedoch weiter wenig zu erkennen. So wurde die Abstimmung von Bund und Ländern in jüngster Vergangenheit u. a. von der Hochschulrektorenkonferenz (2016) und der Deutsche Initiative für Netzwerkinformation (2017) angemahnt – zwei Gremien, in denen

<sup>28</sup> Dokumentation unter: „EUDAT User documentation,“ zuletzt geprüft am 31.01.2017, <https://eudat.eu/services/userdoc/b2find>.

<sup>29</sup> Stand: 06.02.2017

<sup>30</sup> Siehe: „CLARIN,“ zuletzt geprüft am 31.01.2017, <http://www.clarin.eu>.

<sup>31</sup> Siehe: „Global Biodiversity Information Facility,“ zuletzt geprüft am 31.01.2017, <http://www.gbif.org>.

<sup>32</sup> „Bekanntmachung,“ zuletzt geprüft am 31.01.2017, <https://www.bmbf.de/foerderungen/bekanntmachung-1233.html>.

auch Peter Schirnbacher gewirkt hat.

Festzuhalten ist, dass das Zusammenspiel der Informationsinfrastruktureinrichtungen und ihrer Angebote noch am Anfang steht. Für die Zukunft gilt es die Standardisierung der Forschungsdatenrepositorien zu fördern und Richt- und Leitlinien, die sich an Forschende richten, zu harmonisieren. Der Abstimmung der beteiligten Akteure, ob Förderorganisationen, Infrastruktureinrichtungen oder weitere Serviceeinrichtungen, kommt hierbei eine große Bedeutung zu. Dabei gilt es auch die Souveränität der öffentlichen Informationsinfrastruktur zu stärken. Hierzu müssen, wie vom RfII angemahnt, neben Forschungsdateninfrastrukturen auch die Fördermechanismen nachhaltig gestaltet werden. Diese ist besonders wichtig, um die Abhängigkeiten von kommerziell agierenden Infrastrukturen wie Figshare<sup>33</sup> oder Mendeley Data<sup>34</sup> zu vermeiden. Hierzu sind, wie von der HRK und von DINI betont, auch Finanzmittel von Bund und Ländern nötig (Hochschulrektorenkonferenz (2016); Deutsche Initiative für Netzwerkinformation (2017)). Während diese Diskussion auf politischer Ebene zu führen ist, zeigt der Blick in die Praxis bereits jetzt die Erfolge, die im Rahmen der wissenschaftlichen Selbstorganisation entstanden sind. So wurden z. B. unter Mitwirkung von Peter Schirnbacher wichtige Plattformen wie das Wiki Forschungsdaten.org<sup>35</sup> oder der gemeinsamen Arbeitsgruppe Forschungsdaten von DINI und dem Kompetenznetzwerk Nestor geschaffen. Darüber hinaus hat die Bearbeitung des Themenfeldes von Peter Schirnbacher an der Humboldt-Universität zu Berlin mit der Schaffung der Initiative Forschungsdatenmanagement<sup>36</sup> Impulse gegeben, die auch auf andere Hochschulen im In- und Ausland gewirkt haben (Simukovic et al. (2012) ; Kindling et al. (2013)). Diese Arbeit gilt es als Dienstleistung für die Wissenschaft serviceorientiert fortzusetzen und weiter zu professionalisieren. Nicht zuletzt bleibt darüber hinaus zu hoffen, dass dieses wichtige Lehr- und Forschungsgebiet der Informationswissenschaft, das in hohem Maße an der Entwicklung und Etablierung adäquater Informationsinfrastrukturangebote in wissenschaftlichen Einrichtungen sowie an der Vernetzung der Akteure und Stärkung der Community mitgewirkt hat, an der Humboldt-Universität auch nach der Emeritierung von Peter Schirnbacher als Professor fortgeführt wird.

---

<sup>33</sup> Siehe: „figshare,“ zuletzt geprüft am 31.01.2017, <https://figshare.com/mendeley> .

<sup>34</sup> Siehe: „Mendeley Data,“ zuletzt geprüft am 31.01.2017, <https://data.mendeley.com>.

<sup>35</sup> Siehe: „forschungsdaten.org“, zuletzt geprüft am 06.02.2017, <http://www.forschungsdaten.org/>

<sup>36</sup> Siehe: „Forschungsdatenmanagement“, zuletzt geprüft am 06.02.2017, <https://www.cms.hu-berlin.de/dataman>

## Literatur

- Benson, D. A., Clark, K., Karsch-mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. (2015), ‘Dennis A. Benson, Karen Clark, Ilene Karsch-Mizrachi, David J. Lipman, James Ostell and Eric W. Sayers \*’, **43**(March), 30–35.  
**URL:** <https://doi.org/10.1093/nar/gku1216>
- Bloom, T., Ganley, E. & Winker, M. (2014), ‘Data Access for the Open Access Literature: PLOS’s Data Policy’, *PLoS Biology* **12**(2), e1001797.  
**URL:** <https://doi.org/10.1371/journal.pbio.1001797>
- Borgman, C. L. (2007), *Scholarship in the Digital Age. Information, Infrastructure, and the Internet*, MIT Press, Cambridge, Massachusetts.
- Burton, A., Aryani, A., Koers, H., Manghi, P., La Bruzzo, S., Stocker, M., Diepenbroek, M., Schindler, U. & Fenner, M. (2017), ‘The Scholix Framework for Interoperability in Data-Literature Information Exchange’, *D-Lib Magazine* **23**(1/2).  
**URL:** <http://www.dlib.org/dlib/january17/burton/01burton.html>
- Conze, R., Lorenz, H., Ulbricht, D., Elger, K. & Gorgas, T. (2017), ‘Utilizing the International Geo Sample Number Concept in Continental Scientific Drilling During ICDP Expedition COSC-1’, *Data Science Journal* **16**(1), 1–8.  
**URL:** <https://doi.org/10.5334/dsj-2017-002>
- Cousijn, H., Kenall, A., Ganley, E., Harrison, M., Kernohan, D., Murphy, F., Polischuk, P., Martone, M. & Clark, T. (2017), ‘A Data Citation Roadmap for Scientific Publishers’, *bioRxiv*.  
**URL:** <https://doi.org/10.1101/100784>
- Cruse, P., Haak, L. & Pentz, E. (2016), Organization Identifier Project: A Way Forward, Technical report. DOI: 10.5438/2906.  
**URL:** <https://doi.org/10.5438/2906>
- Deutsche Forschungsgemeinschaft (2013), Sicherung guter wissenschaftlicher Praxis, in ‘Sicherung Guter Wissenschaftlicher Praxis’, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, pp. 1–109. DOI: 10.1002/9783527679188.oth1.  
**URL:** <https://doi.org/10.1002/9783527679188.oth1>
- Deutsche Forschungsgemeinschaft (2014), Leitfaden für die Antragstellung. Projektanträge. DFG-Vordruck 54.01 – 06/14, Technical report.
- Deutsche Forschungsgemeinschaft (2015), Leitlinien zum Umgang mit Forschungsdaten, Technical report.  
**URL:** [http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien\\_forschungsdaten.pdf](http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf)

Deutsche Initiative für Netzwerkinformation (2017), Koordiniertes Förderprogramm von Bund und Ländern zum Forschungsdatenmanagement nötig, Technical report.

**URL:** [http://dini.de/service/nachrichten/nachricht/x/\\_/dini\\_hat\\_eine\\_stellungnah/](http://dini.de/service/nachrichten/nachricht/x/_/dini_hat_eine_stellungnah/).

Diepenbroek, M., Grobe, H., Reinke, M., Schindler, U., Schlitzer, R., Sieger, R. & Wefer, G. (2002), 'PANGAEA – an information system for environmental sciences', *Computers & Geosciences* **28**(10), 1201–1210.

European Commission (2012), Online survey on scientific information in the digital age, Technical report, Publications Office of the European Union, Luxembourg.

**URL:** [http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/survey-on-scientific-information-digital-age\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-scientific-information-digital-age_en.pdf)

European Commission (2016), *Open Innovation, Open Science, Open to the World*, number 22 June, Publications Office of the European Union, Luxembourg. DOI: 10.2777/061652.

**URL:** <https://doi.org/10.2777/061652>

Fenner, M., Crosas, M., Grethe, J., Kennedy, D., Hermjakob, H., Rocca-Serra, P., Berjon, R., Karcher, S., Martone, M. & Clark, T. (2016), 'A Data Citation Roadmap for Scholarly Data Repositories', *bioRxiv*.

**URL:** <https://doi.org/10.1101/097196>

Gemeinsamen Wissenschaftskonferenz (2013), Ergebnisse der Sitzung der Gemeinsamen Wissenschaftskonferenz (GWK) am 22. November 2013, Technical report.

**URL:** <http://www.gwk-bonn.de/fileadmin/Pressemitteilungen/pm2013-16.pdf>

Helmholtz-Gemeinschaft (2016), Digitale Forschungsdaten offen zugänglich machen, Technical report.

**URL:** [https://www.helmholtz.de/aktuell/presseinformationen/artikel/artikeldetail/digitale\\_forschungsdaten\\_offen\\_zugaenglich\\_machen/](https://www.helmholtz.de/aktuell/presseinformationen/artikel/artikeldetail/digitale_forschungsdaten_offen_zugaenglich_machen/)

High Level Expert Group on the European Open Science Cloud (2016), Realising the European Open Science Cloud, Technical report.

**URL:** [http://ec.europa.eu/research/openscience/pdf/realising\\_the\\_european\\_open\\_science\\_cloud\\_2016.pdf](http://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf)

Hochschulrektorenkonferenz (2014), 'Management von Forschungsdaten als strategische Aufgabe der Hochschulleitungen'.

**URL:** [https://www.hrk.de/fileadmin/redaktion/hrk/news/\\_migrated/downloads/HRK\\_PM\\_Forschungsdaten\\_14052014.pdf](https://www.hrk.de/fileadmin/redaktion/hrk/news/_migrated/downloads/HRK_PM_Forschungsdaten_14052014.pdf)

Hochschulrektorenkonferenz (2016), Hochschulen sind wichtige Akteure innerhalb der Infrastrukturen für das Forschungsdatenmanagement. Spürbare Impulse von Bund und Ländern sind

unverzichtbar, Technical report.

**URL:** <https://www.hrk.de/themen/forschung/forschungsdaten-management/>

Humboldt-Universität zu Berlin (2014), Grundsätze zum Umgang mit Forschungsdaten an der Humboldt-Universität zu Berlin, Technical report.

**URL:** <https://www.cms.hu-berlin.de/de/dl/dataman/policy>

Kindling, M. & Schirmbacher, P. (2013), ‘„Die digitale Forschungswelt“ als Gegenstand der Forschung / Research on Digital Research / Recherche dans la domaine de la recherche numérique’, *Information - Wissenschaft & Praxis* **64**(2-3).

**URL:** <https://doi.org/10.1515/iwp-2013-0017>

Kindling, M., Schirmbacher, P. & Simukovic, E. (2013), ‘Forschungsdatenmanagement an Hochschulen: Das Beispiel der Humboldt-Universität zu Berlin’, *Libreas* **23**, 43–63.

**URL:** <http://nbn-resolving.de/urn:nbn:de:kobv:11-100212700>

Klump, J., Bertelmann, R., Brase, J., Diepenbroek, M., Grobe, H., Höck, H. & Wächter, J. (2006), ‘Data publication in the open access initiative’, *Data Science Journal* **5**, 79–83.

**URL:** <https://doi.org/10.2481/dsj.5.79>

Kroes, N. (2012), Opening Science Through e-Infrastructures. SPEECH/12/258, Technical report.

**URL:** <http://commentneelie.eu/speech.php?sp=SPEECH/12/258>

Kuipers, T. & Van der Hoeven, J. (2009), Insight into digital preservation of research output in Europe. Survey Report, Technical report, PARSE.Insight.

**URL:** <http://libereurope.eu/wp-content/uploads/2010/01/PARSE.Insight.-Deliverable-D3.4-Survey-Report.-of-research-output-Europe-Title-of-Deliverable-Survey-Report.pdf>

Lehnert, K. A., Klump, J., Arko, R. A., Bristol, S., Buczkowski, B., Chan, C., Chan, S., Conze, R., Cox, S. J. D., Habermann, T., Hangsterfer, A., Hsu, L., Milan, A., Miller, S. P., Noren, A. J., Richard, S. M., Valentine, D. W., Whitenack, T., Wyborn, L. A. & Zaslavsky, I. (2011), IGSN e.V.: Registration and Identification Services for Physical Samples in the Digital Universe, in ‘AGU Fall Meeting 2011’, pp. IN13B–1324.

**URL:** <http://abstractsearch.agu.org/meetings/2011/FM/IN13B-1324.html>

Ludwig, J. & Enke, H., eds (2013), *Leitfaden zum Forschungsdaten-Management. Handreichungen aus dem WissGrid-Projekt*, Verlag Werner Hülsbusch, Glückstadt.

**URL:** [http://www.wissgrid.de/publikationen/Leitfaden\\_Data-Management-WissGrid.pdf](http://www.wissgrid.de/publikationen/Leitfaden_Data-Management-WissGrid.pdf)

Lynch, C. (2003), ‘Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age’, *ARL: A Bimonthly Report* **226**.

**URL:** <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>



Max-Planck-Gesellschaft (2003), Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, Technical report.

**URL:** <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung/>

Nature (2016), ‘Announcement: Where are the data?’, *Nature* **537**(7619), 138–138.

**URL:** <https://doi.org/10.1038/537138a>

Neumann, J. & Brase, J. (2014), ‘DataCite and DOI names for research data’, *Journal of Computer-Aided Molecular Design* **28**(10), 1035–1041.

**URL:** <https://doi.org/10.1007/s10822-014-9776-5>

Pampel, H. & Bertelmann, R. (2011), ‘Data Policies“ im Spannungsfeld zwischen Empfehlung und Verpflichtung, in S. Büttner, H.-C. Hobohm & L. Müller, eds, ‘Handbuch Forschungsdatenmanagement’, Bock + Herchen, Bad Honnef, pp. 49–61.

**URL:** <http://nbn-resolving.de/urn:nbn:de:kobv:525-opus-2287>

Pampel, H. & Dallmeier-Tiessen, S. (2014), Open Research Data: From Vision to Practice, in S. Bartling & S. Friesike, eds, ‘Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing’, Springer, Heidelberg, pp. 213–224. DOI: 10.1007/978-3-319-00026-8\_14.

**URL:** [https://doi.org/10.1007/978-3-319-00026-8\\_14](https://doi.org/10.1007/978-3-319-00026-8_14)

Pampel, H. & Fenner, M. (2016), ‘ORCID - Offener Standard zur Vernetzung von Forschenden’, *Nachrichten aus der Chemie* **64**(1), 57–58.

**URL:** <https://doi.org/10.1002/nadc.20164042239>

Pampel, H., Reuter, E., Scholze, F., Semrau, A., Fritze, F., van de Sandt, S., Goebelbecker, H.-J., Schnepf, E., Schirmbacher, P., Kloska, G., Vierkant, P., Kindling, M., Ulrich, R., Neuberger, N., Bertelmann, R. & Rücknagel, J. (2016), Forschungsdaten-Repositoryen mit Beteiligung deutscher Institutionen. Eine re3data.org-Analyse, Technical report. DOI: 10.5281/zenodo.50148.

**URL:** <https://doi.org/10.5281/zenodo.50148>

Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., Goebelbecker, H.-J., Gundlach, J., Schirmbacher, P. & Dierolf, U. (2013), ‘Making Research Data Repositories Visible: The re3data.org Registry’, *PLOS ONE* **8**(11), e78080.

**URL:** <https://doi.org/10.1371/journal.pone.0078080>

Pfeiffenberger, H. & Carlson, D. (2011), ‘Earth System Science Data’ (ESSD) - A Peer Reviewed Journal for Publication of Data, Vol. 17.

**URL:** <https://doi.org/10.1045/january2011-pfeiffenberger>

Rat für Informationsinfrastrukturen (2016), Leistung aus Vielfalt, Technical report.

**URL:** <http://www.rfi.de/?wpdmdl=1998>

Rücknagel, J., Vierkant, P., Ulrich, R., Kloska, G., Schnepf, E., Fichtmüller, D., Reuter, E., Semrau, A., Kindling, M., Pampel, H., Witt, M., Fritze, F., Van de Sandt, S., Klump, J., Goebelbecker, H.-J., Skarupianski, M., Bertelmann, R., Schirmbacher, P., Scholze, F., Kramer, C., Fuchs, C., Spier, S. & Kirchhoff, A. (2015), Metadata Schema for the Description of Research Data Repositories. Version 3.0, Technical report. DOI: 10.2312/re3.008.

**URL:** <https://doi.org/10.2312/re3.008>

Schirmbacher, P. (2015), 'Informationsmanagement als Bestandteil der bibliotheks- und informationswissenschaftlichen Ausbildung', *Bibliothek Forschung und Praxis* **39**(2), 143–149.

Schäfer, A., Pampel, H., Pfeiffenberger, H., Dallmeier-Tiessen, S., Tissari, S., Darby, R., Giaretta, K., Giaretta, D., Gitmans, K., Helin, H., Lambert, S., Mele, S., Reilly, S., Ruiz, S., Sandberg, M., Schallier, W., Schrimpf, S., Smit, E., Wilkinson, M. & Wilson, M. (2011), Baseline Report on Drivers and Barriers in Data Sharing, Technical report. DOI: 10.5281/zenodo.8306.

**URL:** <https://doi.org/10.5281/zenodo.8306>

Simukovic, E., Kindling, M. & Schirmbacher, P. (2012), 'Unveiling Research Data Stocks : A Case of Humboldt-Universität zu Berlin', *iConference 2014 Proceedings* pp. 742–748.

**URL:** <https://doi.org/10.9776/14351>

Suber, P. (2012), *Open Access*, The MIT Press, Cambridge, Massachusetts.

**URL:** <http://mitpress.mit.edu/books/open-access>

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M. & Frame, M. (2011), 'Data Sharing by Scientists: Practices and Perceptions', *PLoS ONE* **6**(6), e21101.

**URL:** <https://doi.org/10.1371/journal.pone.0021101>

Tenopir, C., Dalton, E. D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D. & Dorsett, K. (2015), 'Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide', *PLOS ONE* **10**(8), e0134826.

**URL:** <https://doi.org/10.1371/journal.pone.0134826>

Ulbricht, D., Elger, K., Bertelmann, R. & Klump, J. (2016), 'panMetaDocs, eSciDoc, and DOIDB—An Infrastructure for the Curation and Publication of File-Based Datasets for GFZ Data Services', *ISPRS International Journal of Geo-Information* **5**(3), 25.

**URL:** <https://doi.org/10.3390/ijgi5030025>

van der Graaf, M. & Waaijers, L. (2011), 'A Surfboard for Riding the Wave. Towards a four country action programme on research data'.

**URL:** [http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=/Files/Filer/downloads/Primary+Research+Data/Surfboard+for+Riding+the+Wave/KE\\_Surfboard\\_Riding\\_the\\_Wave\\_Screen.pdf](http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=/Files/Filer/downloads/Primary+Research+Data/Surfboard+for+Riding+the+Wave/KE_Surfboard_Riding_the_Wave_Screen.pdf)

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., Roche, D. G., Kruuk, L. E. B., Lanfear, R., Binning, S. A., Bechhofer, S., Benson, D. A., Berman, H., Henrick, K., Nakamura, H., Wenger, M., Crosas, M., White, H. C., Carrier, S., Thompson, A., Greenberg, J., Scherle, R., Lecarpentier, D., Martone, M. E., White, E., Sandve, G. K., Nekrutenko, A., Taylor, J., Hovig, E., Wolstencroft, K., Bauch, A., Sansone, S.-A., González-Beltrán, A., Maguire, E., Sansone, S.-A., Rocca-Serra, P., González-Beltrán, A., Harland, L., Groth, P., Berman, H. M., Bourne, P. E., Berman, H. M., Watenpaugh, K., Westbrook, J. D., Fitzgerald, P. M. D., Rose, P. W., Kinjo, A. R., Gutmanas, A., Starr, J. & Musen, M. A. (2016), 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data* **3**, 160018.

**URL:** <https://doi.org/10.1038/sdata.2016.18>

Woutersen-Windhouwer, S. & Brandsma, R. (2009), Enhanced Publications, State of the Art, in M. Verwooy-Gerritsen, ed., 'Enhanced Publications. Linking Publications and Research Data in Digital Repositories', Amsterdam University Press, Amsterdam.

**URL:** <http://hdl.handle.net/11245/1.311760>

Zentralbibliothek für Wirtschaftswissenschaften (2016), 'GeRDI wird Modell für vernetzte Forschungsdaten-Infrastruktur'.

**URL:** <http://www.zbw.eu/de/ueber-uns/aktuelles/meldung/news/gerdi-wird-modell-fuer-vernetzte-forschungsdaten-infrastruktur/>