

Begriffsklärungen

Bericht des Redaktionsausschusses Begriffe an den RfII

Juni 2016

IMPRESSUM

JUNI 2016

Rat für Informationsinfrastrukturen (RfII)
Geschäftsstelle
Papendiek 16
37073 Göttingen

Tel. +49 551 392 09 59
E-Mail info@rfii.de
Web www.rfii.de

MITGLIEDER DES REDAKTIONSAUSSCHUSS BEGRIFFE

Prof. Dr. Thomas Bürger (Leitung), Prof. Dr. Lars Bernard, Sabine Brünger-Weilandt, Prof. Dr. Petra Gehring,
Prof. Dr. Michael Jäckel, Prof. Dr. Klaus Tochtermann

GESTALTUNG

Sarah Baruffaldi

ZITIERVORSCHLAG

RfII – Rat für Informationsinfrastrukturen: Begriffsklärungen. RfII Berichte No. 1, Göttingen 2016, 28 S.

Der RfII bevorzugt eine gendergerechte Sprache. Die in Einzelfällen gewählte männliche Form bezieht gleichermaßen weibliche Personen ein. Auf durchgängige Doppelbezeichnungen wurde aufgrund besserer Lesbarkeit verzichtet.

Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung – Weitergabe unter gleichen Bedingungen 4.0 International Lizenz.



Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

INHALTSVERZEICHNIS

1	Einleitung und Zusammenfassung	1
2	Bisherige Ergebnisse – Zentrale Begriffe	2
2.1	Big Data	2
2.2	Communities, Fächer, Disziplinen	2
2.3	Creative Commons (Lizenzen), Scientific Commons.....	3
2.4	Daten.....	4
2.5	Datenfusion, Datenföderation	4
2.6	Datenlebenszyklus.....	5
2.7	Datenkuration, Datenkuratierung	6
2.8	Datenqualität	6
2.9	Datenschutz, Datensicherheit	7
2.10	Digitale Spaltung	8
2.11	Digitalisierung, Retrodigitalisierung	9
2.12	Dynamische Wissensintegration	10
2.13	Forschungsdaten, Forschungsdatenmanagement	11
2.14	Forschungsformen	12
2.15	Forschungsinfrastrukturen	12
2.16	Informationen	13
2.17	Informationsinfrastrukturen	13
2.18	Informationskompetenz.....	14
2.19	Integration, Interoperabilität	14
2.20	Long Tail Data.....	15
2.21	Medien	16
2.22	Metadaten	16
2.23	Open Access, Open Data – Data Sharing, Open Source	17
2.24	Standards, Standardisierung	18
2.25	Zugänglichkeit, Verfügbarkeit	19
3	Empfehlungen – Begriffsverwendung	20
	Literatur und Onlineressourcen.....	21
	Index.....	27

1 EINLEITUNG UND ZUSAMMENFASSUNG

Der Rat für Informationsinfrastrukturen (RfII) will sich in seinen Empfehlungen einer verständlichen Sprache bedienen. Er strebt dazu einen möglichst einheitlichen Gebrauch häufig und unterschiedlich verwendeter Begriffe an.

Der Redaktionsausschuss Begriffe hat eine Auswahl von Begriffen zusammengestellt, die zunächst für das gewählte Schwerpunktthema Forschungsdaten relevant sind. Die Verwendung in Konzept- und Positionspapieren der Wissenschaftspolitik (v. a. EU, Bund, Länder, Förderer, Fachgemeinschaften etc.) ist berücksichtigt worden, die Begriffe selbst werden möglichst neutral kontextualisiert. Schlagwörter werden vermieden bzw. nur in Ausnahmefällen (z. B. Big Data) berücksichtigt.

Die vorgelegten Begriffsklärungen zielen nicht auf umfassende Definitionen, sondern wollen Orientierungshilfen für Formulierungen des RfII sein. Die Liste der Begriffe ist als lebendiges Dokument konzipiert, und sie wird, auch im Hinblick auf Änderungen im Sprachgebrauch, modifiziert bzw. erweitert werden. Die vorliegende Zusammenstellung erfolgte in den Arbeitsperioden April bis Oktober 2015 sowie Januar bis April 2016.

2 BISHERIGE ERGEBNISSE – ZENTRALE BEGRIFFE

Hiermit werden 25 Begriffsklärungen vorgelegt. Diese sind in folgende Elemente gegliedert:

- Begriff
- englisches Äquivalent (nicht Übersetzbare soll vermieden werden)
- Begriffsklärungsvorschlag (schattiert)
- Verweise: Hinweis auf verwandte Begriffe
- Quellen, Positionspapiere: Hinweise auf Literatur und ggf. Zitate (Auswahl)

2.1 BIG DATA

[big data]

Big Data ist ein häufig verwendetes Schlagwort zur Umschreibung ‚sehr großer‘ komplexer, unstrukturierter und sich schnell ändernder Datenmassen aus heterogenen Datenquellen, für deren datengeleitete Verarbeitung weder die traditionellen Datenanalyseverfahren noch die bestehenden Regeln eines rechtssicheren und ethischen Umgangs mit Daten ausreichen. Der Gebrauch des Schlagworts ist unsicher, da es hinsichtlich der Quantität unbestimmt bleibt: Wo beginnt ‚Big Data‘? Auch wird es häufig mit Intransparenz der Verarbeitung, Nutzung und Verwertung massenhafter Daten zum Nachteil von Persönlichkeitsrechten in Zusammenhang gebracht. Dies führt dazu, dass die Fachwissenschaften das Schlagwort zunehmend meiden.

- Verweise
 - Daten → Datenschutz/Datensicherheit → Forschungsdaten, Forschungsdatenmanagement
 - Zugänglichkeit, Verfügbarkeit
- Quellen, Positionspapiere
 - BDV (2015) – European Big Data Value; GRDI2020 Consortium (2012) – Final Roadmap Report; McKinsey Global Institute (2011) – Big data; Schnell (2013) – Big Data; World Economic Forum (2012) – Big Data.

2.2 COMMUNITIES, FÄCHER, DISZIPLINEN

[communities, disciplines]

Als ‚Community‘ wird – auch im deutschen Sprachgebrauch – eine Gruppe (Gemeinschaft, Gemeinde) von Forschenden bezeichnet, die sozial gut vernetzt ist und ähnlichen Handlungsregeln folgt. Communities können sich um ein gemeinsames Gegenstandsgebiet herum gruppieren (z. B. ‚Klima-Community‘, ‚Heidegger-Community‘), aber auch um Methoden (z. B. ‚HPC-Community‘) oder sogar um Thesen („neurokritische Community“). Der Ausdruck wird nicht selten locker verwendet, Forschende können etlichen Communities angehören, Communities können ggf. rasch entstehen oder sich auch wieder auflösen.

‚Disziplinen‘ (oder traditioneller, mit einem Akzent auf Aspekte der Lehre: ‚Fächer‘) sind demgegenüber die umfassendere Einheit, die Ordnung der Disziplinen (etwa an Hochschulen oder in Ge-

stalt der Fachreferate der Deutschen Forschungsgemeinschaft) ist taxonomischer Art. Vor diesem Hintergrund kann dann wiederum ‚Interdisziplinarität‘ – als gesonderter Modus – explizit zum Problem und Thema werden.

Zur Unterscheidung von ‚Fach‘ und ‚Disziplin‘, wie auch zur umstrittenen Frage nach den Minimalbedingungen, die eine ‚Disziplin‘ auszeichnen (etwa Fachverbände und Fachzeitschriften, ein anerkanntes System von Lehrstuhldenominationen, die Selbstreproduktion durch eigene Studiengänge etc.), gibt es in der Wissenschaftsforschung reichlich Literatur.

Im Zusammenhang von Anforderungen an Informationsinfrastrukturen ist es nicht leicht zu sagen, ob Maßnahmen/Dienste/Angebote besser die Handlungsebene der (kleineren, temporären, relativ homogenen) Communities ansprechen sollten oder aber diejenige der (übergreifenden, in sich heterogenen) Disziplinen und Fächer. Der zum Beispiel vom Wissenschaftsrat nicht selten verwendete Ausdruck ‚Fachgemeinschaft‘ (seltener: ‚Fachcommunity‘) umfasst daher in absichtsvoll unbestimmter Weise beides.

- Verweise
→ Creative Commons (Lizenzen), Scientific Commons → Forschungsformen

2.3 CREATIVE COMMONS (LIZENZEN), SCIENTIFIC COMMONS

[creative commons (licenses) / scientific commons]

Der Begriff Creative Commons (Lizenzierung) bezeichnet etablierte Formen der freiwilligen Kennzeichnung von Informationen und Objekten als kreatives Gemeingut im Internet durch Personen oder Institutionen. Insbesondere in der Wissenschaft werden entsprechend gekennzeichnete Arbeitsergebnisse als Teil der Scientific Commons, des wissenschaftlichen Gemeingutes, verstanden. Das ‚Commons‘-Leitbild steht für größtmögliche Offenheit von Daten, Informationen und Wissen. Die Lizenzkennzeichnung erleichtert einen niedrighwelligen Zugang zu Werken bei gleichzeitiger Wahrung von Rechten des Urhebers. Zwischen auf verbindliches Recht verweisenden Kennzeichnungen proprietärer Inhalte wie „All rights reserved“ [©] und denen der freien Verfügbarkeit bei urheberrechtsfreien Werken (CC0) gibt es mehrere auf freiwilliger Basis akzeptierte Abstufungen mit diesen Kennzeichnungen: CC BY (Urheber muss genannt werden), CC BY-SA (*share alike*, Weitergabe unter gleichen Bedingungen); weitere Kombinationen gibt es mit den Kennzeichnungen -NC (*non-commercial*) und -ND (*no derivatives*).

- Verweise
→ Open Access, Open Data – Data Sharing, Open Source → Zugänglichkeit, Verfügbarkeit
- Quellen, Positionspapiere
Allianz der Wissenschaftsorganisationen (2003) – Berliner Erklärung; Creative Commons – Lizenzen (Webseite).

2.4 DATEN

[data]

Daten sind Grundlage und Teil eines wissenschaftliche Ergebnisse generierenden sowie zugleich ökonomisch wertschöpfenden Kreislaufes aus Daten, Informationen und Wissen. Informationen können aus strukturierten Daten, Wissen aus überprüften Informationen gewonnen werden. Aus wissenschaftlicher Perspektive dienen Daten der Gewinnung und Validierung von (gesichertem) Wissen. Aus ökonomischer Perspektive dienen sie der (wirtschaftlichen) Wertschöpfung. Aus gesellschaftlicher Perspektive sind Daten bildungsrelevant und auch für politische Zwecke vielfach nutzbar.

Grundsätzlich umfasst der Datenbegriff auch nicht-elektronisch bzw. -digital gespeicherte Erhebungen, Berechnungen, Messungen, Zählungen, Texte ebenso wie die Inhalte historischer Überlieferung. Gleichwohl wird der Ausdruck ‚Daten‘ meist im Sinne von ‚digitale Daten‘ verwendet.

Wissenschaftsrelevante Daten entstehen auch außerhalb der Wissenschaften, z. B. statistische Daten aus Verwaltungen, Mobilitäts- oder Telekommunikationsdaten. Der Anteil ungenutzter, verborgener oder wenig erschlossener Daten ist hoch.

Der Wert von Daten hängt entscheidend von der überprüfbaren Datenqualität ab. Dabei können durch eine Anonymisierung von Daten Ansprüche des Datenschutzes und Interessen wissenschaftlicher Auswertung ausgeglichen werden.

Die Verarbeitungs- und Präsentationsmöglichkeiten digitaler Daten revolutionieren und prägen die Informations- und Wissensgesellschaft und erfordern deshalb geeignete öffentlich zugängliche und öffentlich kontrollierte Informationsinfrastrukturen.

- Verweise
→ Forschungsdaten, Forschungsdatenmanagement → Informationen → Informationsinfrastrukturen

2.5 DATENFUSION, DATENFÖDERATION

[data fusion, data federation]

Datenfusion bezeichnet das Zusammenführen von verschiedenen Daten(-sätzen) aus unterschiedlichen Quellen, die alle dasselbe Realwelt-Objekt oder Artefakt beschreiben. Beispielhaft kann das das Zusammenführen unterschiedlicher Beobachtungsdaten zu einem speziellen Umweltphänomen (z. B. einem Erdbeben) oder unterschiedlicher Beschreibungen eines konkreten Kunstwerks sein. Dabei werden die Begriffe Datenfusion und Datenintegration meist synonym verwandt. Die Datenfusion kann unterschiedliche Ziele haben und beispielsweise der Informationsanreicherung zu einem konkreten Objekt, der Schaffung konsistenterer Beschreibungen bei widersprüchlichen Datenlagen oder der Detektion von Änderungen bei Verwendung multitemporaler Datensätze dienen. Die Datenfusion bündelt in der Regel unterschiedliche Verfahren zur Ableitung von Metadaten und zur Datenkuration, zur Harmonisierung der verschiedenen Quelldaten (*mapping*), zur Identifikation übereinstimmender Realwelt-Objekte oder Artefakte in den unterschiedlichen Datensätzen (*matching*) sowie der Ableitung entsprechender Datenverknüpfungen (*links*) und Übereinstimmungswahrscheinlichkeiten.

Die Schaffung einer einheitlichen Beschreibung der verschiedenen Datensätze in den unterschiedlichen Datensammlungen wird in diesem Kontext auch als Datenvirtualisierung bezeichnet, die dazu dient, eine Datenföderation verteilter Datensammlungen derart aufzubauen, dass sich *links* und *matches* schneller auffinden lassen.

- Verweise
 - Datenkuration, Datenkuratierung → Datenqualität → Dynamische Wissensintegration
 - Integration, Interoperabilität → Metadaten
- Quellen, Positionspapiere
 - Datenföderation am konkreten Beispiel: DARIAH-DE – Data Federation Architecture (Webseite); Fachjournal: z. B. Herrera (Hg.) – Information Fusion (Webseite).

2.6 DATENLEBENSZYKLUS

[data life cycle]

Der Datenlebenszyklus ist ein Modell, das den zyklischen Charakter der Arbeit mit Daten aller Art, inkl. Informationen, in ihren verschiedenen Bearbeitungs- und Nutzungsstadien im Prozess der wissenschaftlichen Wertschöpfung beschreibt. Als wesentliche Stadien dieses Zyklus gelten die Datengenerierung (z. B. Messungen), die Datenaufbereitung, die Datenauswertung/-analyse, die Speicherung bis hin zur Langzeitarchivierung sowie die Verfügbarmachung durch Veröffentlichung (z. B. in Datenbanken und Repositorien, als Journal-Publikationen, auf Online-Plattformen) bis hin zur Nachnutzung in weiteren oder neuen Forschungskontexten, die sich auch durch die Lehre ergeben können.

Der zyklische Charakter des Modells hebt hervor, dass durch die Datennutzung und Nachnutzung wiederum neue Ergebnisse in Form von Forschungsdaten generiert werden. Das Datenmanagement entlang dieses Lebenszyklus muss demzufolge sicherstellen, dass Ergebnisse über alle Stadien hinweg reproduzierbar sind. Darüber hinaus sind im Datenlebenszyklus mehrfach Entscheidungen darüber zu treffen, welche Daten aufbewahrt, als Datensatz eigenständig publiziert oder in eine Publikation eingehen werden, und wie lange sie verfügbar zu halten sind, bzw. ob und wann sie gelöscht werden können. Diese Entscheidungen werden aktuell z. B. von Forscherteams oder Einzel Forschern, welche die von ihnen selbst generierten Daten verwalten, nach unterschiedlichen Maßstäben getroffen. Zur Klärung von bestmöglichen organisatorischen Verfahren bis hin zur Langzeitarchivierung besteht Handlungsbedarf, z. B. durch die Entwicklung von Relevanzkriterien und Szenarien.

- Verweise
 - Datensicherheit → Digitalisierung, Retrodigitalisierung → Forschungsdaten, Forschungsdatenmanagement
- Quellen, Positionspapiere
 - KII - Kommission Zukunft der Informationsinfrastruktur (2011) – Gesamtkonzept, S. 17; Datenlebenszyklusmodell: vgl. UK Data Archive (2015) – Research Data Lifecycle (Webseite).

2.7 DATENKURATION, DATENKURATIERUNG

[data curation]

Kuration bzw. Kuratierung bezeichnet im klassischen Sinn die Auswahl und Dokumentation von Elementen einer Sammlung. Datenkuration bezieht sich auf ein erweitertes Aufgabenspektrum, das z. B. die Datenentstehung und Datentransformation in sog. *reichen Metadaten* abbildet. Datenkuration dient der Findbarkeit, Verständlichkeit und Nutzbarkeit von Datensätzen, die in Infrastrukturen aufbewahrt werden. Es handelt sich im Rahmen des Datenmanagements um eine langfristige Service-Aufgabe, die zusätzlich zur ‚Basisdienstleistung‘ der bloßen Speicherung anfällt (Langzeitarchivierung). Diese Service-Aufgabe beinhaltet die kontinuierliche Pflege der Metadaten sowie das Sicherstellen der inhaltlichen und fachlichen Nachnutzbarkeit von Daten. Angesichts der steigenden Datenintensität von Wissenschaft gewinnt die Kuration als Voraussetzung der Nachnutzung rapide an Bedeutung.

Infrastrukturen verfügen idealerweise über Kuratorinnen und Kuratoren, die die Nutzerfreundlichkeit/-nähe von Tools sicherstellen sowie die Aufbereitung von Rohdaten zu wissenschaftlich relevanter Information leisten. Das Aufgabenspektrum der Kuratierung von Daten ist einerseits von technischen Standards anhängig, andererseits von den Nutzeranforderungen. Sie erfordert deshalb hoch qualifiziertes Personal mit fachlichen sowie informationstechnischen Kenntnissen, um z. B. Interoperabilität zwischen Datensätzen in einem Repositorium zu gewährleisten. Die bloße Ablage unkuratierter Daten durch Forschende leistet dies nicht und reduziert das Wertschöpfungspotenzial im Datenlebenszyklus.

- Verweise
→ Datenlebenszyklus → Forschungsdaten, Forschungsdatenmanagement → dynamische Wissensintegration → Metadaten → Zugänglichkeit, Verfügbarkeit
- Quellen, Positionspapiere
Field et al. (2013) – Potential of research data; ICPSR (2013) – Sustaining Domain Repositories; IEAG (2014) – A world that counts; Pempe (2012) – Geisteswissenschaften; WR (2011) – Übergreifende Empfehlungen zu Informationsinfrastrukturen.

2.8 DATENQUALITÄT

[data quality]

Der Begriff Datenqualität bezeichnet die Güte und Verlässlichkeit von Datenobjekten selbst. Voraussetzung für deren wissenschaftliche Nutzbarkeit ist, dass die enthaltenen Daten mit dokumentierten Standards (z. B. von Messgeräten, Berechnungsalgorithmen) und anerkannten Methoden gewonnen werden, und dass diese Standards transparent sind und Nachhaltigkeit sichern. Wenn immer möglich sollte die Unsicherheit in einem Datum geeignet quantifiziert werden.

Die Bewertung von Datenqualität richtet sich nach den zu definierenden Ansprüchen, etwa an die Genauigkeit von Messwerten, abhängig von der jeweiligen Forschungsfrage und damit vom Verwendungszweck. Zudem wird Datenqualität dadurch bestimmt, dass Datensätze und -sammlungen ausreichend Informationen (in Form von möglichst standardisierten technischen und fachlichen Metadaten) über ihre Generierung, Weiterverarbeitung und Analyseverfahren beinhalten. Dies ist

eine wesentliche Voraussetzung für eine Nachnutzung der Daten und die Reproduzierbarkeit von Forschungsergebnissen.

Die abhängig von der Anwendung benötigte Vollständigkeit und Aktualität von Daten und ihren Zusatzinformationen sowie ihre langfristige Verfügbarkeit und Zitierbarkeit sind wiederum Voraussetzungen für die Qualität von Informationsinfrastrukturen und -services, welche die sichere Speicherung, das zielgenaue Auffinden (*Retrieval*), den Zugriff auf die Daten und ihre Nachnutzung (auch im Kontext der Langzeitarchivierung) ermöglichen.

- Verweise
→ Forschungsinfrastrukturen → Informationsinfrastrukturen → Metadaten
- Quellen, Positionspapiere
Zu guter wissenschaftliche Praxis: Allianz der Wissenschaftsorganisationen (2003) – Berliner Erklärung und DFG (2013) – Sicherung guter wissenschaftlicher Praxis, S. 21–22; zu Datenqualität: OECD (2007) – Access to Research Data.

2.9 DATENSCHUTZ, DATENSICHERHEIT

[data protection, data security]

Der Datenschutz, d. h. das Recht auf informationelle Selbstbestimmung, wird aus dem allgemeinen Persönlichkeitsrecht aus Art. 2 Abs. 1 in Verbindung mit Art. 1 Abs. 1 GG hergeleitet. Damit wird die Befugnis des Einzelnen, grundsätzlich selbst über die Preisgabe und Verwendung seiner Daten zu bestimmen, verfassungsrechtlich gewährleistet. Das Datenschutzrecht soll den Einzelnen davor schützen, dass er durch den Umgang mit seinen personenbezogenen Daten beeinträchtigt wird. Im internationalen Sprachgebrauch wird Datenschutz hingegen vielfach mit der US-amerikanischen Rechtsfigur der Privatheit (*Privacy*) verbunden.

Datensicherheit ist ein engerer Begriff als Datenschutz und betrifft Maßnahmen zu dessen technischer und organisatorischer Gewährleistung. Technische Verfahren müssen laut Gesetz so gestaltet werden, dass Vertraulichkeit, Integrität, Verfügbarkeit, Nicht-Verkettbarkeit, Transparenz und Intervenierbarkeit sichergestellt werden. Je komplexer IT-Anwendungen gestaltet sind, desto wichtiger sind effiziente Sicherheitsvorkehrungen. Der Begriff Datensicherheit wird auch im allgemeineren Sinne von IT-Sicherheit verwendet und subsumiert dann alle Aspekte technischer Netzsicherheit, Langzeitarchivierung und -verfügbarkeit sowie die Integrität der Datenhaltung.

Wünsche nach Open Data, nach einem möglichst freien Zugang und einer freien Nachnutzung von Daten für die Wissenschaft, aber auch zugunsten wirtschaftlicher Wertschöpfung, stellen für den Datenschutz wie auch für die Datensicherheit grundsätzlich eine Herausforderung dar. Erwartungen und Forderungen nach Offenheit variieren stark nach Datenarten und Disziplinzugehörigkeit und bedürfen einer differenzierten Betrachtung.

- Verweise
→ Forschungsdaten, Forschungsdatenmanagement → Zugänglichkeit, Verfügbarkeit
- Quellen, Positionspapiere
Zu Datenschutz vs. Open Data: GRDI2020 Consortium (2012) – Final Roadmap Report, S. 97–98;
zu Datenschutz vs. Wertschöpfung: McKinsey Global Institute (2011) – Big data, S. 13; zur Disziplinabhängigkeit: Wagner (2014) – Mehr als nur gute Statistikdaten, S. 7.

2.10 DIGITALE SPALTUNG

[digital divide]

Das Schlagwort ‚digitale Spaltung‘ (digital divide) oder ‚digitale Ungleichheit‘ (digital inequality) steht in der Tradition der Forschungen zur sogenannten ‚Wissenskluft‘ (knowledge gap) und beschreibt die erwarteten gesellschaftlichen Folgen einer zunehmenden Informationsverbreitung durch (Massen-) Medien und neue Technologien. Ursprünglich konnte diese Forschungstradition ihre Popularität aus der Widerlegung der Annahme ableiten, dass eine Steigerung des Informationszuflusses in ein soziales System zu einer allgemeinen Erhöhung des Informations- bzw. Wissensniveaus beiträgt. Die unscharfe Verwendung beider Begriffe – Spaltung und Ungleichheit – wurde dabei immer wieder kritisiert. Unterscheidungen wie „knowledge of“ und „knowledge about“ sowie Fakten- und Strukturwissen ergänzten den allgemeinen Befund, dass das Informationsbedürfnis ungleich verteilt ist und daher „The more, the more“-Effekte eintreten: Ohne Informationszugang sinkt auch der Drang nach Information. Mit dem Aufkommen und der zunehmenden Verbreitung des Internets und mobiler Technologien erweiterte sich das Wirkungsspektrum der digitalen Spaltung, zunächst auf den Zugang zum Internet und zu IT-Technik, dann vermehrt auf qualitative Nutzungsdifferenzen, die sich im Umgang mit Suchmaschinen und anderen Applikationen, aber auch in thematischen Präferenzen (Information vs. Unterhaltung, instrumentelle vs. expressive Nutzung) niederschlugen.

Die Auswirkungen, die der Begriff zusammenfassend anspricht, beziehen sich somit auf die Informationsauswahl, -verarbeitung und -verwertung. Die umfassende Präsenz digitaler Dienste mündet dabei in einen Gewinner-/Verlierer-Diskurs, weil digitale Zugänge und Kompetenzen Voraussetzungen für Inanspruchnahme und Mitwirkung sind. So setzen heute viele Formen der Partizipation die Nutzung von Internettechnologien voraus (Stichwort: „Mitmach-Medien“).

Im wissenschaftlichen Kontext ist zunächst erwartbar, dass die durch Bildungsungleichheit entstehenden Differenzen geringer ausfallen. Gleichwohl lassen sich auch im wissenschaftlichen Bereich Unterschiede identifizieren, die aus dem Vorhandensein neuer Informationsinfrastrukturen hervorgehen, und zwar im Hinblick auf:

- die Arbeitsweise und Recherche;
 - die Informationskompetenz, insbesondere der Forschenden und Studierenden (auch mit Blick auf unterschiedliche Disziplinen);
 - die Neuausrichtung didaktischer Konzepte;
 - den Zugang zu digitalen Angeboten (Datenbanken, Portale);
 - die Zugänglichkeit lizenzierter Inhalte;
 - Open-Access-Policies;
 - neue Kollaborationsformen.
-
- Verweise
 - Informationskompetenz → Zugänglichkeit, Verfügbarkeit
-
- Quellen, Positionspapiere
 - Allianz-Initiative Digitale Information - AG Forschungsdaten (2015) – Research data at your fingertips, S. 3–4; HRK (2015) – Handlungsoptionen für Forschungsdatenmanagement, S. 15; Hyman/Sheatsley (1947) – Information Campaigns.

2.11 DIGITALISIERUNG, RETRODIGITALISIERUNG

[digitization]

Digitalisierung bezeichnet ganz allgemein die Umstellung der gesamten Gesellschaft auf die Verwendung von Digitaltechnologien (digitale Revolution, *Digital Turn*).

Im Zusammenhang wissenschaftlicher Informationsinfrastrukturen umschreibt Digitalisierung die Umstrukturierung von Studien- und Forschungsprozessen mit digitalen Methoden und Werkzeugen: Kommunikation über Social Media; E-Learning / Blended Learning; kollaboratives Arbeiten in virtuellen Forschungsumgebungen; elektronisches Publizieren etc.

Die digitale Transformation von analoger Schrift-, Bild- oder Tonüberlieferung, von Sammlungen und Objekten wird auch als Retrodigitalisierung bezeichnet. Sie umfasst den Vorgang von der Metadaten-Erschließung über das Scannen bis zur Speicherung und digitalen Langzeitarchivierung. Digitalisierte ‚Images‘ von Texten und Bildern werden mit Texterkennungsprogrammen (z. B. Optical Character Recognition – OCR) zu maschinenlesbaren Volltexten und mit Bilderkennungsverfahren zu durchsuchbaren Bilddatenbanken aufbereitet. Damit werden Text- und Bildkorpora neu erschlossen und stehen für innovative Analysemethoden (text mining u. v. a.) zur Verfügung.

- Verweise
 - Daten → Datenlebenszyklus → Informationen → Informationsinfrastrukturen → Open Access, Open Data – Data Sharing, Open Source
- Quellen, Positionspapiere
 - DFG (2012) – Digitale Transformation.

2.12 DYNAMISCHE WISSENSINTEGRATION

[dynamic knowledge management/knowledge integration]

Die in Deutschland bisher kaum gebräuchliche Wendung ‚dynamische Wissensintegration‘ wird vom Rfll als bildlicher Begriff für die neue Welt des komplexen Managements von Forschungsdaten genutzt. Er greift auf Formulierungen wie *knowledge integration* oder *knowledge integration dynamics* zurück, wie sie seit den 1980er Jahren in zahlreichen Studien zum Organisations- und Wissensmanagement insbesondere der Forschungs- und Entwicklungsabteilungen großer Industrieunternehmen verwendet werden.

Die bewusst unscharfe Formulierung umschreibt das Ziel, Daten, Informationen und Wissen von Forschern und Forschergruppen zur Lösung komplexer Aufgaben über mehrere Forschungszyklen zusammenzuführen, um Innovationsprozesse und Synergieentwicklungen in neuer Qualität zu ermöglichen. In diesem zyklischen Prozess entstehen aus Forschungsdaten (-materialien) neue Ergebnisse (Produkte), die wiederum zu Forschungsdaten in einer nächsten Evaluationsphase werden. Der Erkenntnisgewinn über mehrere Zyklen hängt damit zusammen, wie gut es gelingt, diese Dynamik der Wissensgewinnung nachvollziehbar zu dokumentieren und überprüfbar zu halten. Dies ist besonders relevant, wenn sich in späteren Zyklen zeigt, dass Messungen oder Interpretationen aus früheren Zyklen falsch sind oder seinerzeit unbekanntem Einflussfaktoren unterliegen. Viele Forschungs- und Entwicklungsfragen lassen sich nur in Zusammenarbeit unterschiedlicher Disziplinen, Fächer und Communities durch eine dynamische Integration verteilter Wissensressourcen über die Zeit effizient lösen.

Dazu bedarf es nicht nur der Bereitschaft und Fähigkeit enger Kommunikation und Kooperation in internen und externen Netzwerken, sondern auch geeigneter Werkzeuge, Verfahren und Infrastrukturen. Diese müssen von Wissenschaftlern, Datenspezialisten (Data Scientists) und geeigneten Institutionen gemeinsam ausgebaut und zu leistungsstarken, forschungsnahen und nachhaltigen Informationsinfrastrukturen bedarfsgerecht entwickelt werden.

Interoperabilität und Integrierbarkeit sollen sicherstellen, dass die in Forschungsprozessen gewonnenen Ergebnisse (Forschungsdaten, Publikationen, Projektdokumentationen) in auf Nachhaltigkeit, Zugänglichkeit und Verfügbarkeit ausgerichteten Informationsinfrastrukturen nach transparenten internationalen Standards gespeichert, dokumentiert, erschlossen und zur weiteren Wertschöpfung anschlussfähig aufbereitet und angeboten werden.

- Verweise
→ Datenlebenszyklus → Forschungsinfrastrukturen → Informationsinfrastrukturen → Integration, Interoperabilität
- Quellen, Positionspapiere
Haddad/Bozdogan (2009) – Knowledge Integration; Nováček et al. (2008) – Infrastructure; Wahlstedt (2014) – Dynamic Knowledge Integration.

2.13 FORSCHUNGSDATEN, FORSCHUNGSDATENMANAGEMENT

[research data, research data management]

Forschungsdaten sind Daten, die im Zuge wissenschaftlicher Vorhaben entstehen, z. B. durch Beobachtungen, Experimente, Simulationsrechnungen, Erhebungen, Befragungen, Quellenforschungen, Aufzeichnungen, Digitalisierung, Auswertungen. Forschungspragmatisch, wenn auch nicht immer trennscharf, lassen sich Forschungsprimärdaten von -sekundärdaten unterscheiden, die den Entstehungsprozess der Primärdaten dokumentieren und kontextualisieren. Im Forschungsprozess können Sekundärdaten selbst wieder zu Primärdaten werden, was für den Lebenszyklus von Forschungsdaten von Bedeutung ist.

Das Forschungsdatenmanagement umfasst alle – über das Forscherhandeln im engeren Sinne hinaus auch organisationsbezogenen – Maßnahmen, die getroffen werden müssen, um qualitätsvolle Daten zu gewinnen, um die gute wissenschaftliche Praxis im Datenlebenszyklus einzuhalten, um Ergebnisse reproduzierbar und Daten zur Nachnutzung verfügbar zu machen und um ggf. bestehenden Dokumentationsverpflichtungen (z. B. im Gesundheitswesen) Rechnung zu tragen. Zunehmend fordern Förderorganisationen von Projektnehmern die Erstellung eines Datenmanagementplans sowie von wissenschaftlichen Institutionen, dass sie sich auf die Erfordernisse eines Nachhaltigkeit gewährleistenden Datenmanagements einstellen. Datenmanagementpläne zu Anfang eines Projekts oder einer Forschungsarbeit sind geeignet, die zu nutzenden und zu generierenden Daten und die notwendigen Dokumentationen, Metadaten und Standards zu beschreiben, mögliche rechtliche Einschränkungen (z. B. Datenschutz) rechtzeitig zu benennen, benötigte Speicherressourcen einzuplanen sowie Kriterien festzulegen, welche Daten Externen in welcher Form verfügbar gemacht werden und wie langfristig die Daten zu sichern sind. Auf der Organisationsebene müssen forschende Einrichtungen (z. B. Hochschulen) den Zugang zu entsprechenden Infrastrukturdiensten innerhalb der Einrichtung (z. B. durch Auf- und Ausbau passender Kapazitäten) oder in Zusammenarbeit mit externen Partnern (durch Kooperationsverträge etc.) sichern.

- Verweise
→ Daten → Datenlebenszyklus → Forschungsinfrastrukturen → Metadaten
- Quellen, Positionspapiere
Allianz-Initiative Digitale Information - AG Forschungsdaten (2015) – Research data at your fingertips; zu Forschungsdaten = Grundlage von Argumentation und Rechnen: European Commission - DG for Research and Innovation (2013) – Guidelines on Open Access, S. 3; zu Forschungsdaten = Primärquelle wissenschaftlicher Aktivität: OECD (2007) – Access to Research Data, S. 13; zu Forschungsdaten aus Perspektive der Sozialwissenschaften: RatSWD (2010) – Kriterien Forschungsdaten-Infrastruktur, S. 4; zu Forschungsdaten als Daten aus dem Forschungsprozess: Allianz-Initiative Digitale Information (2012) – Leitbild, S. 7; WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 53–57; DCC - Digital Curation Center (2015) – Data Management Plans (Webseite); DFG (2015) – Leitlinien Forschungsdaten; HRK (2014) – Management von Forschungsdaten; HRK (2012) – Hochschule im digitalen Zeitalter.

2.14 FORSCHUNGSFORMEN

[modes of research]

„Zur Beschreibung der Perspektiven für die Weiterentwicklung von Informationsinfrastrukturen unterscheidet der Wissenschaftsrat typologisch sechs fächerübergreifende Forschungsformen: experimentierende, beobachtende, hermeneutisch-interpretierende, begrifflich-theoretische und gestaltende Forschungsformen sowie Simulationen. Mit diesen Forschungsformen verbinden sich jeweils spezifische Anforderungen an Informationsinfrastrukturen.“ (WR 2012, S. 8).

- Verweise
→ Daten → Forschungsdaten, Forschungsdatenmanagement → Informationen → Medien
- Quellen, Positionspapiere
WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 8, 25-40; WR (2014) – Positionspapier Simulation.

2.15 FORSCHUNGSINFRASTRUKTUREN

[research infrastructures]

Forschungsinfrastrukturen sind der Forschung dienende wissenschaftliche Infrastrukturen (Anlagen, Ressourcen, Einrichtungen und Dienstleistungen). Darunter fallen:

- a. Großgeräte oder Instrumente für Forschungszwecke (z. B. Forschungsschiffe, Satelliten- und Raumfahrtstationen, Teleskope, Teilchenbeschleuniger),
- b. Informations- und Wissensressourcen wie (nicht-digitale oder digitale) Sammlungen, Archive, Bibliotheken, Datenbanken,
- c. informations- und kommunikationstechnische Infrastrukturen wie Rechner, Rechnernetze (GRID, Cloud),
- d. Software sowie
- e. jegliche sonstige für die wissenschaftliche Forschung in vergleichbarer Funktion genutzte Anlage, Ressource, Einrichtung oder Dienstleistung.

- Verweise
→ Forschungsdaten, Forschungsdatenmanagement → Informationsinfrastrukturen
- Quellen, Positionspapiere
Forschungsinfrastrukturen = langlebige Investitionsgüter für Wissenschaft: BMBF (2013) – Pilotprojekt nationale Roadmap, S. 2; Forschungsinfrastrukturen = langlebige, offene Ressourcen für die Wissenschaft: BMBF (2015) – Leitfaden Nationale Roadmap, S. 5; Umfassende Definition von Forschungsinfrastrukturen: European Commission (2014) – Horizon 2020 Work Programme 2014 und ESFRI (2011) – Strategy Report, S. 7; Forschungsinfrastrukturen = Instrumente, Ressourcen oder Services von strategischer Bedeutung: WR (2013) – Bewertung Forschungsinfrastrukturvorhaben.

2.16 INFORMATIONEN

[information]

Informationen (im Plural) sind – informationswissenschaftlich gesehen – allgemeine Voraussetzungen für Kenntnisse, in der Verwertung aber ergebnisoffen. Sie können verarbeitet, kontextualisiert oder interpretiert werden. Im Prozess der Wissensgewinnung stehen Informationen idealtypisch zwischen Daten und Wissen. Informationsinfrastrukturen haben diesbezüglich eine Dienstleisterrolle. Für die wissenschaftliche Wissensgewinnung und die ökonomische Wertschöpfung spielt die Arbeit mit Informationen eine zentrale Rolle. Darüber hinaus sind Fragen von Qualität, Validität und Neuheit (bzw. sogar Einzigartigkeit) von Informationen mit besonderem wissenschaftlichen Gewicht oder ökonomischen Wert.

- Verweise
→ Daten → Datenqualität → Forschungsdaten, Forschungsdatenmanagement → Informationsinfrastrukturen → Zugänglichkeit, Verfügbarkeit

2.17 INFORMATIONENINFRASTRUKTUREN

[information infrastructures (Rfll), e-infrastructures (EU)]

Informationsinfrastrukturen sind technisch und organisatorisch vernetzte Dienste und Angebote für den Zugang zu und die Erhaltung von Daten-, Informations- und Wissensbeständen. Im Sinne des Rfll dienen sie primär Forschungszwecken, sie sind häufig Forschungsgegenstand und haben stets eine ermöglichende Funktion.

Informationsinfrastrukturen müssen berücksichtigen, dass Wissensbestände in Universitäten, Forschungseinrichtungen, Archiven, Bibliotheken und Museen in analogen, digitalen oder in Mischformen vorliegen. Die digitale Erschließung analoger Wissensbestände zielt auf die Integration und Konvergenz zwischen digitalisierten und nativ digitalen Daten in einheitlichen, integrierten Arbeitsumgebungen mit dem Ziel dynamischer Wissensintegration. Wie der englische Ausdruck ‚e-Infrastructures‘ referenziert deshalb auch der deutsche Begriff ‚Informationsinfrastrukturen‘ zunehmend auf digitale Informations- und Kommunikationstechnologien für die Forschung.

Die Leistungsfähigkeit von digitalen Informationsinfrastrukturen hängt maßgeblich von den Investitionen für die Erschließung der Inhalte, nutzungsfreundlichen Zugangsformen, technischer Ausstattung, internationalen Standards und effektiven Werkzeugen ab. Ebenso relevant ist die informationsfachliche Kompetenz von Nutzern und Personal und – damit zusammenhängend – die Qualität passgenauer Dienstleistungen.

- Verweise
→ Daten → Forschungsinfrastrukturen → Informationen
- Quellen, Positionspapiere
Zu e-Infrastructures, insbesondere Hardware und Services: z. B. e-IRG (2013) – White Paper (long), S. 5–10 und JISC - Joint Information Systems Committee – e-Infrastructure Programme (Webseite); zu Informationsinfrastrukturen: z. B. GWK (2013) – Drucksache 13.48 und WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 7, 30.

2.18 INFORMATIONSKOMPETENZ

[digital literacy, information literacy]

Der Begriff Informationskompetenz bezeichnet eine Schlüsselkompetenz für breite Schichten der Informationsgesellschaft. Er subsumiert spezifischere Begriffe wie Medienkompetenz, digitale Kompetenz, Lese-, Schreib- und Rechenkompetenz (im Englischen verschiedene Varianten von *literacy*). Sie wird als Voraussetzung einer umfassenden Kommunikations- und Handlungskompetenz betrachtet und thematisiert den Umgang mit sowie die Verwertung von Informationsmaterialien unterschiedlichen Formates, Umfangs und Bezuges durch Nutzerkreise mit variierenden Qualitätsansprüchen. Dabei lassen sich grob eine instrumentell-professionelle (Schule, Studium, Arbeitswelt) und eine reflektierend-allgemeinbildende Dimension nennen. Der Kompetenzerwerb soll zu einer selbstgesteuerten Nutzung befähigen und eine verantwortungsvolle Be- und Verwertung von Informationen befördern. Informationskompetenz umfasst dabei ausdrücklich auch die Erstellung von Inhalten für Dritte („create messages“). Mit Hilfe der digitalen Informationsinfrastrukturen muss auf allen Ebenen des Bildungssystems die Vermittlung von Informations- und Medienkompetenz systematisch ausgebaut werden.

- Verweise
→ Datenschutz, Datensicherheit → Datenqualität → Zugänglichkeit, Verfügbarkeit
- Quellen, Positionspapiere
DBV (2013) – Stellungnahme zur Informationskompetenz; Euro Media Literacy – European Charter for Media (Webseite); HRK (2012) – Hochschule im digitalen Zeitalter.

2.19 INTEGRATION, INTEROPERABILITÄT

[integration, interoperability]

Seiner lateinischen Wurzel nach hebt das Wort Integration auf die unbeschadete (Wieder-)Herstellung ab, es hat sich aber sowohl im generellen Sprachgebrauch zerredet als auch in Fachsprachen verschiedenste Bedeutungen angenommen. Im Zusammenhang von Informationsinfrastrukturen sowie IT-Systemen, Software und Datenfragen, müssen insbesondere folgende (zuweilen auch ineinander übergehende) Verwendungen des Begriffs unterschieden werden:

Als Integration kann ein Prozess oder eine Verbindungs-/Zusammenführungsleistung bezeichnet werden, bei welcher die verbundenen Elemente intakt oder sogar unverändert bleiben; Integration ist dann Effekt. So können Daten in Softwareumgebungen integriert werden, neue Anforderungen können in etablierte Standards integriert werden, verteiltes Wissen wird zusammengeführt und zu einem neuen Ganzen integriert („integrated knowledge“) etc.

Als Integration können aber auch solche Formen der Einbindung bezeichnet werden, bei welcher die ‚Nahtlosigkeit‘ des Überganges zwischen den verbundenen Elementen im Vordergrund steht; Integration („gute‘ Integration oder auch ‚Integriertheit‘) ist dann tendenziell eine Eigenschaft oder ein Zustand. So können komplexe IT-Systeme ‚hoch integriert‘ sein, eine Landschaft von Forschungsinfrastrukturen kann durch einen hohen Integrationsgrad gekennzeichnet sein etc.

Ein Aspekt von Integration in diesem Sinne ist die ‚Interoperabilität‘. Dieser Begriff zielt auf den pragmatisch entscheidenden Punkt ab, dass Geräte oder Systeme mit (ggf. noch unbekanntem) Fremdgeräten oder Fremdsystemen kommunizieren/zusammenarbeiten können – auch in flexiblen Szenarien. Als technische Voraussetzung für Interoperabilität gelten transparente sowie hinreichend standardisierte Schnittstellen.

Schließlich kann Integration als Prozess betrachtet werden, der sich an Außengrenzen abspielt (und dabei u. a. das Maß an Offenheit/Aufnahmefähigkeit des durch sie begrenzten Ganzen charakterisiert). In diesem Sinne wird sowohl in IT-Zusammenhängen als auch bezogen auf Organisationen, Kulturen etc. von ‚Integrationsfähigkeit‘ oder auf Integration angelegten Strukturen gesprochen.

- Verweise
→ Datenfusion, Datenföderation → dynamische Wissensintegration → Metadaten → Standards, Standardisierung
- Quellen, Positionspapiere
Brünger-Weilandt (2014) – Informationsinfrastruktur; Gradman (2015) – Interoperability; Ludwig/Enke (Hg.) (2013) – Leitfaden Forschungsdatenmanagement; van de Sompel/Nelson (2015) – 15 Years of Interoperability Efforts.

2.20 LONG TAIL DATA

[long tail data]

Der Begriff long tail leitet sich aus der Darstellung von Verteilungskurven für unterschiedliche Aspekte des Datenmanagements und der Datennutzung ab. Solche Verteilungen können für eine große Menge von Datensätzen meist nur – je nach Art der betrachteten abhängigen Variable (Y-Achse) – geringe Einnahmen, Nutzungszahlen, Datengrößen, Beschreibungstiefen, Harmonisierungsgrade usw. aufzeigen und weisen damit einen langen Schwanz (long tail) in der entsprechenden Verteilungskurve auf. Diese Long-tail-Daten stehen daher oft nicht zuerst im Fokus der Datenanbieter und Datennutzer. Dennoch können auch Long-tail-Daten wertvolle Grundlage für die wissenschaftliche Arbeit sein.

So wird z. B. mit der Digitalisierung analoger Sammlungen von Texten, Bildern und Objekten ein Long-tail-Effekt verbunden, wie er aus der Internetökonomie bekannt ist. Wie schlecht verkäufliche Lagerware erst im Internet auf lohnende Nachfrage stößt, so können bislang unausgeschöpfte Wissensbestände durch Digitalisierung weltweit bekannt werden. Die Nutzung von Long-tail-Daten trägt dann zur wissenschaftlichen und wirtschaftlichen Wertschöpfung bei.

- Verweise
→ Daten → Datenlebenszyklus → Digitalisierung, Retrodigitalisierung → Informationsstrukturen
- Quellen, Positionspapiere
Expert Advisory Group for Research Infrastructures (2014) – Consultation Horizon 2020; Heidorn (2008) – Shedding Light on the Dark; HRK (2015) – Handlungsoptionen für Forschungsdatenmanagement, S. 12; RDA – Long tail of research data Interest Group (Webseite); Wallis et al. (2013) – If We Share Data.

2.21 MEDIEN

[media (units)]

Medien (im Plural) sind im pragmatischen Sprachgebrauch der Informations- und Bibliothekswissenschaften in integrierender Weise Träger und Vermittler/Verbreiter von (und zwar jeweils zugleich) Daten, Informationen und Wissen. Als Datenträger haben sich vielfältige analoge und digitale Formate etabliert, die jedoch nicht alle eine aktive Verbreitungsfunktion im Sinne von Massenmedien (z. B. Buch, Zeitung, Radio, Fernsehen, Internet) übernehmen. Medien ermöglichen das Verfügbarhalten von Inhalten und/oder das Erreichen von Adressaten. Das jeweilige digitale oder analoge Format bestimmt den Zugang zu den Inhalten. Informationsdienste stellen somit stets medientypisch *formatgebundene* Information zur Verfügung.

Medien vermitteln, sichern und verbreiten Primär- und Sekundärdaten und können im Forschungsprozess selbst zum Untersuchungsgegenstand, zur Datenquelle und damit zu Primärdaten werden.

- Verweise
→ Daten → Forschungsformen → Zugänglichkeit, Verfügbarkeit
- Quellen, Positionspapiere
Medienbegriff = weit gefasst, inkludiert natürliche Objekte etc.: WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 16–17.

2.22 METADATEN

[meta data]

Metadaten sind Daten über Daten. Sie dienen der Beschreibung von Daten und unterstützen Datennutzer bei der Recherche nach Daten, bei der Bewertung der Eignung recherchierter Daten für die eigenen Zwecke und bei der Integration gefundener Daten in die eigene Systemumgebung. Metadaten decken dabei typischerweise die Beschreibung syntaktischer (z. B. Formatangaben), semantischer (z. B. Angaben zu Dateninhalten) qualitativer (z. B. Angaben zur Genauigkeit bzw. Datenunsicherheit) und rechtlicher Aspekte (z. B. Nutzungsrechte) sowie der Datenherkunft und -vorverarbeitung ab und beschreiben den Zugriffsweg zu den beschriebenen Daten (z. B. eine Internetadresse).

Dabei finden sich Metadaten in ganz unterschiedlichen Strukturierungs- und Harmonisierungsgraden und sind unterschiedlich vollständig. Zur Schaffung und Harmonisierung von Metadatenbeständen liegen zahlreiche internationale, auch disziplinspezifische Standards vor, die ihren Ursprung häufig im Bibliothekswesen haben. Ein bekanntes Beispiel ist der Dublin Core Standard, jüngst ist der internationale Metadatenstandard „Resource Description and Access“ als gemeinsamer Rahmen für die Beschreibung von digitalen und analogen Ressourcen in Bibliotheken, Archiven und Museen eingeführt worden. Die „Metadata Standards Directory Working Group“ der Research Data Alliance unterstützt die Entwicklung, Anwendung und Transparenz von Metadatenstandards für Forschungsdaten und führt ein Verzeichnis.

Verfahren und Entwicklungen zur automatisierten Ableitung (Annotation) von Metadaten aus existierenden und entstehenden digitalen Datenbestände machen Metadaten für die weitere Bearbeitung erschließbar. Dabei können Metadaten auch selbst zu Forschungsdaten werden.

- Verweise
 - Daten → Datenfusion, Datenföderation → Datenqualität → Integration, Interoperabilität
 - Standards, Standardisierung

- Quellen, Positionspapiere
 - Allianz-Initiative Digitale Information (2012) – Leitbild, S. 8, 14; DCMI – Dublin Core Metadata Initiative – Specifications (Webseite); European Commission (2010) – Riding the wave, S. 20; JSC – Joint Steering Committee for Development of RDA – Resource Description and Access (Webseite); KII - Kommission Zukunft der Informationsinfrastruktur (2011) – Gesamtkonzept, S. 50; RDA – Metadata Standards Directory Working Group (Webseite); Metadaten = Dokumentation des Inhalts: Berman (2008) – Got data?, S. 55; Metadaten = Schwelle zwischen Forschung und Infrastruktur: WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 20–22; Metadaten bedarfsgerecht einsetzen: WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 51; Veröffentlichungen umfassen auch Metadaten: Allianz der Wissenschaftsorganisationen (2003) – Berliner Erklärung, S. 2.

2.23 OPEN ACCESS, OPEN DATA – DATA SHARING, OPEN SOURCE

[open access]

Open Access (OA) (oder auch, im Zugriffsanspruch weitergehend: Open Data und Open Content) bezeichnet den niedrigschwelligen, möglichst freien und breiten Zugang zu Ergebnissen, die dem Wissenschaftssystem entstammen. Zahlreiche Forschungsförderer und wesentliche Teile der Forschung selbst präferieren inzwischen Open Access (auch aus der Industrie wird ein solch „offener“ Zugang zu Ergebnissen öffentlicher Forschung gefordert). Zwei Modelle der öffentlichen Ko-Finanzierung von Open-Access-Publikationen haben sich etabliert: der ‚Grüne Weg‘ (Zweitveröffentlichung in digitaler Form) und der ‚Goldene Weg‘ (digitale Erstveröffentlichung mit dem Ziel sofortiger und größtmöglicher Sichtbarkeit und Verfügbarkeit).

In keinem Fall ist unter Open Access eine gänzlich schrankenlose Verwertung zu verstehen. Auch OA-Publikationen unterliegen rechtlichen Regeln (insbesondere dem UrhG). Dies gilt auch für Modelle des auf technischer Ebene realisierten oder anderweitig vereinbarten und praktizierten unentgeltlichen Teilens von Daten (Open Data bzw. Data Sharing) oder Programmen (Open Source). Durch Vergabe differenzierter Creative-Commons-Lizenzen können die Inhaber von Urheber- und Leistungsschutzrechten Interessierten weitere Freiheiten (z. B. der Verarbeitung oder Verwertung) einräumen. Der geforderte Open Access nicht nur zu Publikationen, sondern auch zu Forschungsdaten hat im internationalen Diskurs an Gewicht gewonnen. Ebenso hat sich die bildungspolitische Forderung nach Open Educational Resources international – in Deutschland allerdings vergleichsweise spät – inzwischen breit durchgesetzt.

In der zunehmend durch Digitalisierung geprägten Wissenschaft wird (im Unterschied zur Wirtschaft) der Gemeingutcharakter der Forschungsressourcen, der Zwischen- wie Endergebnisse sowie auch des Teilens von Daten als Chance und Katalysator kollaborativer Wissensgenerierung ge-

schätzt. Durch Digitaltechnik sind rechtliche und wirtschaftliche Rahmenbedingungen, die der Wissenschaft ihre wissenschaftstypischen Freiräume sichern, in Bewegung geraten. Der Zugang zu Forschungsdaten muss deshalb durch erweiterte Schrankenregelungen zu Gunsten von Bildung und Wissenschaft gesichert und organisiert werden, z. B. indem sie von Verwertung freigestellt oder ggf. dem öffentlichen Interesse gemäß entgolten werden. Eine Möglichkeit hierfür ist eine z. B. durch Forschende, Bibliotheken etc. proaktiv vorgenommene Lizenzierung von Daten oder Publikationen als ‚offen‘ nutzbares Gut.

Eine wesentliche Frage ist, ob auf zugänglich gemachte Daten und Inhalte weiter aufgebaut werden kann, ob diese analysiert, kombiniert, oder auch verbessert werden können. In manchen Fällen ist ein reiner Zugang ausreichend, in anderen Fällen, insbesondere im Falle von Open Educational Resources und Forschungsdaten sind Weiternutzungsrechte (etwa in Form von Creative-Commons-Lizenzen) essenziell.

- Verweise
→ Creative Commons (Lizenzen), Scientific Commons → Zugänglichkeit, Verfügbarkeit

- Quellen, Positionspapiere
Zu Open Access: u. a. Allianz der Wissenschaftsorganisationen (2003) – Berliner Erklärung; European Commission (2012) – Access to Scientific Information; zu Open Educational Resources: Bundesregierung (2013) – Koalitionsvertrag; zu verschiedenen Formen des Data Sharing in der Wissenschaft: KE - Knowledge Exchange (2014) – Sowing the seed, S. 12–14.

2.24 STANDARDS, STANDARDISIERUNG

[standards, standardization]

Standards und Regeln dienen bei der Prozessierung von Daten, Information und Wissen dem Ziel der Normierung und damit einer Form der Qualitätssicherung (Qualität im Sinne von Wiederverwendbarkeit) in inhaltlicher, formaler, rechtlicher, organisatorischer und technischer Hinsicht. Schon im frühen klassischen Bibliotheks- und Archivwesen wurden Formate für Informationsmedien definiert und Regeln für ihre Verzeichnung (Metadaten) und Speicherung festgelegt. Das 1917 gegründete Deutsche Institut für Normung e. V. (DIN) organisierte im öffentlichen Interesse Standardisierungen zur Beachtung anerkannter Regeln der Technik in Wirtschaft, Wissenschaft und Verwaltung mit den Zielen der Qualitätssicherung und Gebrauchstauglichkeit von Innovationen. Zu den bekanntesten deutschen Normen zählen die DIN-Formate des Papiers aus dem Jahr 1922. Die International Organization for Standardization (ISO) wurde 1947 in Genf gegründet

Die Digitalisierung aller Bereiche der Wissenschaft und die hohe Dynamik digitaler Technik erfordern Standardisierungen insbesondere bei Daten und Metadaten, Austauschformaten, Schnittstellen, Datenmodellen, Auszeichnungssprachen und Vokabularen. Es gibt eine Vielzahl von Dokumentenformaten wie z. B. PDF oder HTML und eine noch größere Zahl medientypologisch definierter Metadatenformate wie z. B. Dublin Core für Objektbeschreibungen im Internet, EAD für Handschriften und Archivgut oder MARC, MAB, MODS für bibliothekarische Austauschformate etc., jedoch keine allgemein verbindlichen Standardformate für Metadaten. Dies erlaubt Individualisierungsmöglichkeiten für medienspezifische Beschreibungen und erschwert zugleich Interoperabilität und Kompatibilität von Daten und Dokumenten unterschiedlicher Provenienz.

Für Daten, Software-Programme und -Prozesse, für Dokumentationen oder das Prozessmanagement gibt es inzwischen eine ganze Reihe wichtiger ISO-Normen. Anwendung, Einhaltung und koordinierte Weiterentwicklung definierter internationaler Standards und Normen sind Voraussetzungen sowohl für Integrität und Vertrauenswürdigkeit als auch für nachhaltige und effiziente Informationsinfrastrukturen zur Gewährleistung von Gebrauchsfähigkeit, Interoperabilität, Durchlässigkeit, Anschlussfähigkeit und Langzeitverfügbarkeit.

Standardisierungen bzw. die Einhaltung von Standards beim Datenmanagement sind aufwändige, jedoch wissenschaftlich und wirtschaftlich sinnvolle Investitionen in die Qualitätssicherung und Nachhaltigkeit. Wissenschaftliche Informationsinfrastrukturen sollten, wenn immer sinnvoll und geboten, in internationale Standardisierungsprozesse eingebunden sein.

- Verweise
→ Daten → Datenqualität → Integration, Interoperabilität → Metadaten
- Quellen, Positionspapiere
Allianz der Wissenschaftsorganisationen (2010) – Grundsätze Forschungsdaten; DFG (2012) – Digitale Transformation; DSA - Data Seal of Approval (2013) – Guidelines Version 2; Force 11 – FAIR data principles (Webseite); Witt, Kroll et al. (2012) – ISO 16363: Digital Repository Certification; WR (2016) – Kerndatensatz Forschung.

2.25 ZUGÄNGLICHKEIT, VERFÜGBARKEIT

[accessibility, availability, authorized access]

Erfolgreiche und effektive Forschung beruht auf möglichst barrierefreier Zugänglichkeit von Forschungsergebnissen und -ressourcen, also auf ‚offen‘ geregelter, niedrigschwelliger Nutzbarkeit (Open Information, Open Data, Open Educational Resources). Einen offenen Zugang zu wissenschaftlichem Wissen – nicht nur als Bildungs-, sondern auch als Forschungsressource – haben die deutschen Wissenschaftsorganisationen mehrfach gefordert. Wissenschaftler benötigen Informationsinfrastrukturen, die ihnen mit möglichst geringen praktischen, wirtschaftlichen oder rechtlichen Barrieren Zugang zu Daten, Informationen und Wissen zur Verfügung stellen. Ein Schlüssel hierfür sind beispielsweise Modelle des Open-Access-Publizierens, institutionell, regional, national oder auf Communities ausgerichtete Lizenzierungsmodelle, die unter Verzicht auf Entgelt den Zugang zum Zweck nichtwirtschaftlicher (z. B. wissenschaftlicher) Nutzung vorsehen, oder Formen des Teilens von Daten (Data Sharing) und Informationen – bis hin zu ihrer wirtschaftlichen Kollektivierung (Commons).

Die Verfügbarkeit von Daten und Informationen geht über deren technisch, organisatorisch und rechtlich zu lösende Zugänglichkeit hinaus. Verfügbarkeit umfasst alle Aspekte

- a. der Erschließung von Daten, von der Vergabe von Ressourcen-IDs/Bezeichnern/Identifizierungscodes über die Metadaten bis zur differenzierten maschinellen Indexierung und/oder intellektuellen Tiefenerschließung nach fachspezifischen Standards
- b. der Weiterverarbeitung von Im- und Exporten bis hin zur weiteren Verbreitung und Nachnutzbarkeit in neuen Kontexten (Open Science, Open Educational Resources) und

c. der Speicherung bis hin zur befristeten oder unbefristeten Archivierung mit dem Ziel von langfristig uneingeschränktem Zugriff und uneingeschränkter Lesbarkeit durch fortlaufende Sicherung der Inhalte.

- Verweise
→ Creative Commons, Scientific Commons → Datenschutz, Datensicherheit → Metadaten
→ Open Access, Open Data – Data Sharing, Open Source
- Quellen, Positionspapiere
Allianz der Wissenschaftsorganisationen (2003) – Berliner Erklärung; European Commission (2012) – Access to Scientific Information, S. 3; KII - Kommission Zukunft der Informationsinfrastruktur (2011) – Gesamtkonzept; Neuroth et al. (Hg.) (2012) – Langzeitarchivierung, S. 9, 17; RatSWD (2010) – Kriterien Forschungsdaten-Infrastruktur, S. 4; WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 60, 76; zum Einsatz von Identifiern: DINI-AG Forschungsinformationssysteme (2015) – Forschungsinformationssysteme.

3 EMPFEHLUNGEN – BEGRIFFSVERWENDUNG

In seiner Auftakterklärung hat der RfII sich das Ziel gesetzt, „struktur- und prozessbezogene [...] Empfehlungen zu erarbeiten, die dem Bedarf an der langfristig sicheren, standortunabhängigen sowie rechtlich, finanziell und technisch gesicherten langfristigen Verfügbarkeit wissenschaftlich relevanter Daten, Informationen und Wissensbestände Rechnung tragen.“¹ Um dieses Ziel zu erreichen, schien es geboten, den Redaktionsausschuss Begriffe mit grundlegenden Begriffsklärungen zur Bestimmung des sich daraus ergebenden ‚Spielfeldes‘ für den RfII zu beauftragen.

Die Arbeit des Redaktionsausschusses zeigt, dass Begriffsklärungen innerhalb des etablierten, komplexen Infrastrukturdiskurses zwar aufwändig, aber notwendig sind. Gegenwärtig birgt ein von hoher Diversität geprägter Diskurs die Gefahr von Missverständnissen. Der Ausschuss empfiehlt deshalb, die vorliegende Liste geklärter Begriffe als sprachliche Orientierung für kommende Kommunikationsaufgaben zu verwenden. Darüber hinaus sollte es Ziel des RfII sein, im Bereich wissenschaftlicher Infrastrukturen gemeinsam mit anderen Akteuren an der Gestaltung einer verständlichen Sprache mitzuwirken und sparsam mit Wortneuschöpfungen, Trendwörtern und Metaphern umzugehen.

Im Laufe des Jahres 2017 wird der RfII seine Begriffeliste grundlegend überprüfen.

¹ RfII - Rat für Informationsinfrastrukturen (2015) – Auftakterklärung.

LITERATUR UND ONLINERESSOURCEN

- Allianz der Wissenschaftsorganisationen (2003): Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen, Berlin, online verfügbar unter:
http://openaccess.mpg.de/68053/Berliner_Erklaerung_dt_Version_07-2006.pdf, zuletzt geprüft am: 25.04.2016.
- Allianz der Wissenschaftsorganisationen (2010): Grundsätze zum Umgang mit Forschungsdaten, online verfügbar unter:
http://www.allianzinitiative.de/fileadmin/user_upload/redakteur/Grundsätze_Forschungsdaten_2010.pdf, zuletzt geprüft am: 25.04.2016.
- Allianz-Initiative Digitale Information (2012): Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen – Fortsetzung der Zusammenarbeit in den Jahren 2013 bis 2017, München, online verfügbar unter:
http://www.allianzinitiative.de/fileadmin/user_upload/redakteur/Schwerpunktinitiative_2013-2017.pdf, zuletzt geprüft am: 25.04.2016.
- Allianz-Initiative Digitale Information - AG Forschungsdaten (2015): Positionspapier „Research data at your fingertips“. Hg. v. Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen, o.O., DOI: 10.2312/allianzfd.001.
- BDV - Big Data Value Europe (2015): European Big Data Value. Strategic Research & Innovation Agenda. Version 1.0, Brüssel, online verfügbar unter: http://ok-bdva.iais.fraunhofer.de/sites/default/files/europeanbigdatavaluepartnership_sria__v1_0_final.pdf, zuletzt geprüft am: 25.04.2016.
- Berman, Francine (2008): Got data? A Guide to Data Preservation in the Information Age, in: CACM - Communications of the ACM 51, Nr. 12, S. 50–56, DOI: 10.1145/1409360.1409376.
- BMBF - Bundesministerium für Bildung und Forschung (2013): Roadmap für Forschungsinfrastrukturen. Pilotprojekt, Bonn, online verfügbar unter:
http://www.bmbf.de/pub/roadmap_forschungsinfrastrukturen.pdf, zuletzt geprüft am: 25.04.2016.
- BMBF - Bundesministerium für Bildung und Forschung (2015): Leitfaden zur Konzepterstellung für die Nationale Roadmap für Forschungsinfrastrukturen, Bonn, online verfügbar unter:
https://www.bmbf.de/pub/leitfaden_zur_konzepterstellung_forschungsinfrastruktur.pdf, zuletzt geprüft am: 25.04.2016.
- Brünger-Weilandt, Sabine (2014): Informationsinfrastruktur – eine Standortbestimmung, in: Ceynowa, Klaus/Hermann, Martin (Hg.): Bibliotheken: Innovation aus Tradition. Rolf Griebel zum 65. Geburtstag. Festschrift, Berlin: De Gruyter Saur, S. 138–150, online verfügbar unter:
<http://www.degruyter.com/view/books/9783110310511/9783110310511.138/9783110310511.138.xml>, zuletzt geprüft am: 25.04.2016.
- Bundesregierung (2013): Deutschlands Zukunft gestalten. Koalitionsvertrag zwischen CDU, CSU und SPD. 18. Legislaturperiode, Berlin, online verfügbar unter:
<http://www.cdu.de/sites/default/files/media/dokumente/koalitionsvertrag.pdf>, zuletzt geprüft am: 25.04.2016.
- Creative Commons: Lizenzen (Webseite), online verfügbar unter:
<http://creativecommons.org/licenses/>, zuletzt geprüft am: 25.04.2016.
- DARIAH-DE: Data Federation Architecture (Webseite), Göttingen, online verfügbar unter:
<https://de.dariah.eu/data-federation-architecture>, zuletzt geprüft am: 25.04.2016.

- DBV - Deutscher Bibliotheksverband (2013): Stellungnahme des Deutschen Bibliotheksverbands zur Entschließung der 13. Mitgliederversammlung der Hochschulrektorenkonferenz „Hochschule im digitalen Zeitalter: Informationskompetenz neu begreifen – Prozesse anders steuern“, Berlin, online verfügbar unter:
http://www.bibliotheksverband.de/fileadmin/user_upload/DBV/positionen/2013_10_14_Stellungnahme_Informationskompetenz_endg.pdf, zuletzt geprüft am: 25.04.2016.
- DCC - Digital Curation Center (2015): Data Management Plans (Webseite), online verfügbar unter:
<http://www.dcc.ac.uk/resources/data-management-plans>, zuletzt geprüft am: 25.04.2016.
- DCMI – Dublin Core Metadata Initiative: Specifications (Webseite), online verfügbar unter:
<http://dublincore.org/specifications/>, zuletzt geprüft am: 25.04.2016.
- DFG - Deutsche Forschungsgemeinschaft (2012): Die digitale Transformation weiter gestalten. Der Beitrag der Deutschen Forschungsgemeinschaft zu einer innovativen Informationsinfrastruktur für die Forschung. Positionspapier. Hg. v. DFG-Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (AWBI), Bonn, online verfügbar unter:
http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier_digitale_transformation.pdf, zuletzt geprüft am: 25.04.2016.
- DFG - Deutsche Forschungsgemeinschaft (2013): Sicherung guter wissenschaftlicher Praxis. Denkschrift. Empfehlungen der Kommission zur Selbstkontrolle in der Wissenschaft, Bonn, online verfügbar unter:
http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf, zuletzt geprüft am: 25.04.2016.
- DFG - Deutsche Forschungsgemeinschaft (2015): Leitlinien zum Umgang mit Forschungsdaten, Bonn, online verfügbar unter:
http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf, zuletzt geprüft am: 25.04.2016.
- DINI-AG Forschungsinformationssysteme (2015): Forschungsinformationssysteme in Hochschulen und Forschungseinrichtungen. Positionspapier. Version 1.0 (DINI Schriften, 15-de), Göttingen, online verfügbar unter: https://zenodo.org/record/14828/files/AGFIS_Positionspapier.pdf, zuletzt geprüft am: 25.04.2016.
- DSA - Data Seal of Approval (2013): Guidelines Version 2 - 2014/2015, online verfügbar unter:
http://datasealofapproval.org/media/filer_public/2013/09/27/guidelines_2014-2015.pdf, zuletzt geprüft am: 25.04.2016.
- e-IRG - e-Infrastructures Reflection Group (2013): White Paper. Long Version, Den Haag, online verfügbar unter: http://e-irg.eu/documents/10920/11274/annex_5.2_e-irg_white_paper_2013_-_final_version.pdf, zuletzt geprüft am: 25.04.2016.
- ESFRI - European Strategy Forum on Research Infrastructures (2011): Strategy Report on Research Infrastructures. Roadmap 2010, Luxembourg, DOI: 10.2777/23127.
- Euro Media Literacy: The European Charter for Media Literacy (Webseite), online verfügbar unter:
<http://www.euromedialiteracy.eu/charter.php?id=5>, zuletzt geprüft am: 25.04.2016.
- European Commission (2012): Towards better access to scientific information. Boosting the benefits of public investments in research. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2012) 401 final, Brüssel, online verfügbar unter:
<http://ec.europa.eu/transparency/regdoc/rep/1/2012/EN/1-2012-401-EN-F1-1.Pdf>, zuletzt geprüft am: 25.04.2016.

- European Commission (2014): Horizon 2020 Work Programme 2014-2015. Revised. Part 4: European research infrastructures (including e-Infrastructures). European Commission Decision C (2014) 4995 of 22 July 2014, Brüssel, online verfügbar unter: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-infrastructures_en.pdf, zuletzt geprüft am: 25.04.2016.
- European Commission - Directorate-General for Research and Innovation (2013): Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. Version 1.0, Brüssel, online verfügbar unter: http://www.gsrt.gr/EOX/files/h2020-hi-oa-pilot-guide_en.pdf, zuletzt geprüft am: 25.04.2016.
- European Commission (2010): 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, o.O., online verfügbar unter: http://ec.europa.eu/information_society/newsroom/cf/itemlongdetail.cfm?item_id=6204, zuletzt geprüft am: 25.04.2016.
- Expert Advisory Group for Research Infrastructures (2014): Consultation of the Horizon 2020 Advisory Groups. Providing advice on potential priorities for Research and Innovation in the Work Programme 2016-2017. Unter Mitarbeit von Sabine Brünger-Weilandt, o.O., online verfügbar unter: <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetailDoc&id=14950&no=1>, zuletzt geprüft am: 25.04.2016.
- Field, Laurence et al. (2013): Realising the full potential of research data. Common challenges in data management, sharing and integration across scientific disciplines. Version 3, o.O., online verfügbar unter: http://orca.cf.ac.uk/66034/1/ESFRI_Common_Challenges_v1.pdf, zuletzt geprüft am: 25.04.2016.
- Force 11: The fair data guiding principles. For comment (Webseite), online verfügbar unter: <https://www.force11.org/group/fairgroup/fairprinciples>, zuletzt geprüft am: 25.04.2016.
- Gradman, Stefan (2015): Interoperability. A key concept for large scale, persistent digital libraries. Digital Preservation Europe, o.O., online verfügbar unter: <http://www.arcamemorie.net/wp-content/uploads/2015/04/2gradnman-briefing-paper-graphic.pdf>, zuletzt geprüft am: 25.04.2016.
- GRDI2020 Consortium (2012): GRDI2020 Final Roadmap Report. Global Research Data Infrastructures: The Big Data Challenges, o.O., online verfügbar unter: <http://www.grdi2020.eu/Repository/FileScaricati/e2b03611-e58f-4242-946a-5b21f17d2947.pdf>, zuletzt geprüft am: 25.04.2016.
- GWK - Gemeinsame Wissenschaftskonferenz (2013): Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen. Ergänzter Bericht des Ausschusses (GWK Drucksache 13.48, unveröffentlicht), Bonn.
- Haddad, Marc/Bozdogan, Kirkor (2009): Knowledge Integration in Large-Scale Organizations and Networks. Conceptual Overview and Operational Definition, Cambridge (MA), DOI: 10.2139/ssrn.1437029.
- Heidorn, P. Bryan (2008): Shedding Light on the Dark Data in the Long Tail of Science, in: Library Trends 57, Nr. 2, online verfügbar unter: https://www.ideals.illinois.edu/bitstream/handle/2142/9127/Heidorn_LongTail_PreprintwEdits.doc.pdf?sequence=7, zuletzt geprüft am: 25.04.2016.
- Herrera, Francisco (Hg.): Information Fusion. An International Journal on Multi-Sensor, Multi-Source Information Fusion (Webseite), online verfügbar unter: <http://www.journals.elsevier.com/information-fusion/>, zuletzt geprüft am: 25.04.2016.

- HRK - Hochschulrektorenkonferenz (2012): Hochschule im digitalen Zeitalter. Informationskompetenz neu begreifen – Prozesse anders steuern. Entschließung der 13. Mitgliederversammlung der HRK am 20. November 2012 in Göttingen, Bonn, online verfügbar unter: http://www.hrk.de/uploads/media/Entschliessung_Informationskompetenz_20112012_01.pdf, zuletzt geprüft am: 25.04.2016.
- HRK - Hochschulrektorenkonferenz (2014): Management von Forschungsdaten. Eine zentrale strategische Herausforderung für Hochschulleitungen. Empfehlung der 16. Mitgliederversammlung der HRK am 13. Mai 2014 in Frankfurt am Main, Bonn, online verfügbar unter: https://www.hrk.de/uploads/tx_szconvention/HRK_Empfehlung_Forschungsdaten_13052014_01.pdf, zuletzt geprüft am: 25.04.2016.
- HRK - Hochschulrektorenkonferenz (2015): Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können. Orientierungspfade, Handlungsoptionen, Szenarien. Empfehlungen der 19. Mitgliederversammlung, Kiel, online verfügbar unter: http://www.hrk.de/uploads/tx_szconvention/Empfehlung_Forschungsdatenmanagement__final_Stand_11.11.2015.pdf, zuletzt geprüft am: 25.04.2016.
- Hyman, Herbert/Sheatsley, Paul (1947): Some Reasons why Information Campaigns Fail, in: Public Opinion Quarterly 11, S. 412–423, online verfügbar unter: <http://poq.oxfordjournals.org/content/11/3/412.full.pdf>, zuletzt geprüft am: 25.04.2016.
- ICPSR - Interuniversity Consortium for Political and Social Research (2013): Sustaining Domain Repositories for Digital Data. A White Paper, o.O., online verfügbar unter: http://datacommunity.icpsr.umich.edu/sites/default/files/WhitePaper_ICPSR_SDRDD_121113.pdf, zuletzt geprüft am: 25.04.2016.
- IEAG - Independent Expert Advisory Group on a Data Revolution for Sustainable Development (2014): A world that counts. Mobilising the data revolution for sustainable development. United Nations Secretary General, o.O., online verfügbar unter: <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>, zuletzt geprüft am: 25.04.2016.
- JISC - Joint Information Systems Committee: e-Infrastructure Programme (Webseite), Bristol, online verfügbar unter: <http://webarchive.nationalarchives.gov.uk/20140702233839/http://www.jisc.ac.uk/whatwedo/programmes/einfrastructure.aspx>, zuletzt geprüft am: 25.04.2016.
- JSC – Joint Steering Committee for Development of RDA: Resource Description and Access (Webseite), online verfügbar unter: <http://www.rda-jsc.org/archivedsite/rda.html>, zuletzt geprüft am: 25.04.2016.
- KE - Knowledge Exchange (2014): Sowing the seed. Incentives and motivations for sharing research data, a researcher’s perspective, Kopenhagen, online verfügbar unter: http://repository.jisc.ac.uk/5662/1/KE_report-incentives-for-sharing-researchdata.pdf, zuletzt geprüft am: 25.04.2016.
- KII - Kommission Zukunft der Informationsinfrastruktur (2011): Gesamtkonzept für die Informationsinfrastruktur in Deutschland. Empfehlungen im Auftrag der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder. Unter Mitarbeit von Sabine Brünger-Weilandt, o.O., online verfügbar unter: http://www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Infrastruktur/KII_Gesamtkonzept.pdf, zuletzt geprüft am: 25.04.2016.
- Ludwig, Jens/Enke, Harry (Hg.) (2013): Leitfaden zum Forschungsdaten-Management. Handreichungen zum WissGrid-Projekt, Glückstadt/Göttingen: Hülsbusch/Universitätsverlag Göttingen, online verfügbar unter: http://www.wissgrid.de/publikationen/Leitfaden_Data-Management-WissGrid.pdf, zuletzt geprüft am: 25.04.2016.

- McKinsey Global Institute (2011): Big Data. The next frontier for innovation, competition, and productivity, o.O., online verfügbar unter: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation, zuletzt geprüft am: 25.04.2016.
- Neuroth, Heike et al. (Hg.) (2012): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme. D-Grid/Nestor/BMBF - Bundesministerium für Bildung und Forschung, Boizenburg/Göttingen: Hülsbusch/Universitätsverlag Göttingen, online verfügbar unter: <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>, zuletzt geprüft am: 25.04.2016.
- Nováček, Vít et al. (2008): Infrastructure for dynamic knowledge integration. Automated biomedical ontology extension using textual resources, in: *Journal of Biomedical Informatics* 41, Nr. 5, S. 816–828, DOI: 10.1016/j.jbi.2008.06.003.
- OECD - Organisation for economic co-operation and development (2007): OECD Principles and Guidelines for Access to Research Data from Public Funding, Paris, online verfügbar unter: <http://www.oecd.org/sti/sci-tech/38500813.pdf>, zuletzt geprüft am: 25.04.2016.
- Pempe, Wolfgang (2012): Geisteswissenschaften, in: Neuroth, Heike et al. (Hg.): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme, Boizenburg/Göttingen: Hülsbusch/Universitätsverlag Göttingen, S. 137–159, online verfügbar unter: http://nestor.sub.uni-goettingen.de/bestandsaufnahme/kapitel/nestor_bestandsaufnahme_007.pdf, zuletzt geprüft am: 25.04.2016.
- RatSWD - Rat für Sozial- und Wirtschaftsdaten (2010): Kriterien des RatSWD für die Einrichtung der Forschungsdaten-Infrastruktur, Berlin, online verfügbar unter: http://www.ratswd.de/download/publikationen_rat/RatSWD_FDZKriterien.PDF, zuletzt geprüft am: 25.04.2016.
- RDA - Research Data Alliance: Long tail of research data Interest Group (Webseite), online verfügbar unter: <https://rd-alliance.org/groups/long-tail-research-data-ig.html>, zuletzt geprüft am: 25.04.2016.
- RDA - Research Data Alliance: Metadata Standards Directory Working Group (Webseite), online verfügbar unter: <http://rd-alliance.github.io/metadata-directory/>, zuletzt geprüft am: 25.04.2016.
- RfII - Rat für Informationsinfrastrukturen (2015): Auftakterklärung, Göttingen, DOI: 10.5281/zenodo.50662.
- Schnell, Rainer (2013): Getting Big Data but avoiding Big Brother (Working Paper Series, WP-GRLC-2013-02). Hg. v. GRLC - German Record Linkage Center, online verfügbar unter: <http://www.record-linkage.de/-download=wp-grlc-2013-02.pdf>, zuletzt geprüft am: 25.04.2016.
- UK Data Archive (2015): Research Data Lifecycle (Webseite), online verfügbar unter: <http://data-archive.ac.uk/create-manage/life-cycle>, zuletzt geprüft am: 25.04.2016.
- van de Sompel, Herbert/Nelson, Michael L. (2015): Reminiscing about 15 Years of Interoperability Efforts, in: *D-Lib Magazine* 21, Nr. 11/12, DOI: 10.1045/november2015-vandesompel.
- Wagner, Gert (2014): Zu guter Forschungsinfrastruktur und forschungsbasierter Politikberatung gehören mehr als nur gute Statistikdaten (RatSWD Working Paper Series, 231). Hg. v. RatSWD - Rat für Sozial- und Wirtschaftsdaten, Berlin, online verfügbar unter: http://ratswd.de/dl/RatSWD_WP_231.pdf, zuletzt geprüft am: 25.04.2016.
- Wahlstedt, Linnéa (2014): Dynamic Knowledge Integration. A field study of an Information Systems Development Project (Linköping Studies in Arts and Science, 603), Linköping: Linköping University, online verfügbar unter: <http://liu.diva-portal.org/smash/get/diva2:679589/FULLTEXT01.pdf>, zuletzt geprüft am: 25.04.2016.

- Wallis, Jillian C./Rolando, Elizabeth/Borgman, Christine L. (2013): If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology, in: PLoS ONE 8, Nr. 7, S. 1–17, DOI: 10.1371/journal.pone.0067332.
- Witt, Michael et al. (2012): ISO 16363: Trustworthy Digital Repository Certification in Practice (Libraries Faculty and Staff Presentations, 4). Hg. v. Purdue University, West Lafayette, online verfügbar unter: http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1003&context=lib_fspress, zuletzt geprüft am: 25.04.2016.
- World Economic Forum (2012): Big Data, Big Impact. New Possibilities for International Development, Genf, online verfügbar unter: http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf, zuletzt geprüft am: 25.04.2016.
- WR - Wissenschaftsrat (2011): Übergreifende Empfehlungen zu Informationsinfrastrukturen. Drs. 10466-11, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/10466-11.pdf>, zuletzt geprüft am: 25.04.2016.
- WR - Wissenschaftsrat (2012): Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020. Drs. 2359-12, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>, zuletzt geprüft am: 25.04.2016.
- WR - Wissenschaftsrat (2013): Bericht zur wissenschaftsgeleiteten Bewertung umfangreicher Forschungsinfrastrukturvorhaben für die Nationale Roadmap (Pilotphase). Drs. 2841-13, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/2841-13.pdf>, zuletzt geprüft am: 25.04.2016.
- WR - Wissenschaftsrat (2014): Bedeutung und Weiterentwicklung von Simulation in der Wissenschaft. Positionspapier. Drs. 4032-14, Dresden, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/4032-14.pdf>, zuletzt geprüft am: 25.04.2016.
- WR - Wissenschaftsrat (2016): Empfehlungen zur Spezifikation des Kerndatensatz Forschung. Drs. 5066-16, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/5066-16.pdf>, zuletzt geprüft am: 25.04.2016.

INDEX

Big Data	4, 11, 17
Commons	5, 19, 20, 21, 22
Communities	4, 5, 12, 21
Creative Commons	5, 20, 22
Data Sharing	5, 11, 19, 20, 21, 22
Daten	4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23
Datenföderation	6, 7, 17, 19
Datenfusion	6, 17, 19
Datenkuration.....	6, 7, 8
Datenlebenszyklus	7, 8, 11, 12, 13, 17
Datenqualität.....	6, 7, 8, 9, 15, 16, 19, 21
Datenschutz.....	4, 9, 13, 16, 22
Datensicherheit	4, 7, 9
Dienste.....	5, 10, 15
Digital Divide.....	10
Digitale Spaltung.....	10
Digitalisierung.....	7, 11, 13, 17, 19, 20
Disziplinen	4
Dynamische Wissensintegration.....	7, 8, 12, 17
Fächer	4, 5, 12
Forschungsdaten	3, 4, 6, 7, 8, 10, 12, 13, 14, 15, 18, 19, 20
Forschungsdatenmanagement	4, 6, 7, 8, 13, 14, 15, 17
Forschungsformen.....	5, 14, 18
Forschungsinfrastrukturen	9, 12, 13, 14, 15, 16
Informationen.....	5, 6, 7, 8, 11, 12, 14, 15, 16, 18, 20, 21, 23
Informationsinfrastrukturen.....	3, 5, 6, 9, 10, 11, 12, 14, 15, 16, 17, 21
Informationskompetenz	11, 16
Integration.....	7, 12, 15, 16, 17, 18, 19, 21
Interoperabilität	7, 8, 12, 16, 17, 19, 20, 21
Lizenzierung.....	5, 20
Long Tail Data	17
Medien	10, 11, 14, 18
Metadaten.....	6, 7, 8, 9, 11, 13, 17, 18, 19, 20, 21, 22
Nachhaltigkeit.....	8, 12, 13, 21
Nutzbarkeit.....	8, 21
Objekt	5, 6, 17, 18
Open Access	5, 11, 12, 19, 21, 22
Open Data.....	5, 9, 10, 11, 19, 21, 22
Open Science	21
Open Source	5, 11, 19, 22
Ressourcen	14, 18, 21
Retrodigitalisierung	7, 11, 17
Scientific Commons	5, 20
Standardisierung.....	17, 19, 20

Standards.....	8, 12, 13, 15, 16, 17, 18, 19, 20, 21
Verfügbarkeit.....	4, 5, 8, 9, 10, 12, 15, 16, 18, 19, 20, 21, 23
Zugänglichkeit.....	4, 5, 8, 10, 11, 12, 15, 16, 17, 18, 20, 21