



Rat für
I n f o r m a t i o n s
I n f r a s t r u k t u r e n

LEISTUNG AUS VIELFALT

Empfehlungen zu Strukturen, Prozessen und Finanzierung
des Forschungsdatenmanagements in Deutschland

IMPRESSUM

Verabschiedet am 03.05.2016

Rat für Informationsinfrastrukturen (RfII)
Geschäftsstelle
Papendiek 16
37073 Göttingen

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

GESTALTUNG, SATZ UND DRUCK
Basta Werbeagentur GmbH, Göttingen

ZITIERVORSCHLAG

RfII – Rat für Informationsinfrastrukturen: Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland, Göttingen 2016, 160 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.

INHALT

Kurzfassung	1
Executive Summary.....	5
1 Herausforderungen und Aufgabenstellung	9
2 Ausgangslage.....	12
2.1 Forschungsdaten, Informationsinfrastrukturen und Wissenschaftspolitik	12
2.2 Überblick über die „Landschaft“	15
2.3 Trägerschaft – Finanzierung – Förderprogramme.....	18
2.4 Deutschland im internationalen Vergleich	24
2.5 Ermöglichungsstrukturen und besondere Entwicklungsaufgaben ...	31
3 Positionen zu Grundfragen	34
4 Empfehlungen	37
4.1 Verstetigung von Infrastrukturprojekten.....	37
4.2 Schwerpunktbildung und Arbeitsteilung im System	39
4.3 Langzeitarchivierung und Langzeitverfügbarkeit von Forschungsdaten.....	45
4.4 Neu zu etablierende Akteure/Zuständigkeiten	48
4.5 Neue Berufsbilder, Studiengänge, Ausbildungswege.....	49
4.6 Neue „Datenkultur“	52
4.7 Nutzungsanreize und innerwissenschaftliche Akzeptanz	54
4.8 Internationale Forschungs- und Informationsinfrastrukturpolitik	56
4.9 Nutzung und Verwertung	59
4.10 Mehr Wissen über Forschungsdaten	60
4.11 Datenschutz	61
4.12 Datensicherheit.....	62
4.13 Finanzierung und Nachhaltigkeit.....	63
5 Ausblick und Priorisierung der Empfehlungen	66
Literatur	70
Onlineressourcen.....	75
Anhang	78
A. Begriffsklärungen.....	A-1
B. Die Entwicklung von Konzepten für Informationsinfrastrukturen in der Bundesrepublik Deutschland seit den 1960er Jahren.....	B-1
C. Szenarien des Forschungsdatenmanagements	C-1
D. Daten und Fakten.....	D-1
E. Mitwirkende	E-1

Abkürzungsverzeichnis

BMBF	Bundesministerium für Bildung und Forschung
DARA	Registrierungsagentur für Sozial- und Wirtschaftsdaten
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DFG	Deutsche Forschungsgemeinschaft
DFN-Verein	Verein zur Förderung eines Deutschen Forschungsnetzes e.V.
DINI	Deutsche Initiative für Netzwerkinformationen e.V.
DIPF	Deutsches Institut für Internationale Pädagogische Forschung
DMP	Datenmanagementplan
ELIXIR	European Life-Sciences Infrastructure for Biological Information
EOSC	European Open Science Cloud
ERIC	European Research Infrastructure Consortium
ESFRI	European Strategy Forum on Research Infrastructures
EUDAT	European Data Infrastructure
FAIR	Findable, Accessible, Interoperable, Reusable
FDM	Forschungsdatenmanagement
GBIF	Global Biodiversity Information Facility
GESIS	Leibniz Institut für Sozialwissenschaften
GFBio	German Federation for Biological Data
GWDG	Gesellschaft für wissenschaftliche Datenverarbeitung mbH, Göttingen
GWK	Gemeinsame Wissenschaftskonferenz
HPC	High Performance Computing
HR	Human Resources
HRK	Hochschulrektorenkonferenz
IQB	Institut zur Qualitätsentwicklung im Bildungswesen
KII	Kommission Zukunft der Informationsinfrastruktur
KMU	Kleine und mittlere Unternehmen
NEPS	Nationales Bildungspanel (National Education Panel Study)
NFDI	Nationale Forschungsdateninfrastruktur
NHR	Nationales Hoch- und Höchstleistungsrechnen
NoMaD	Novel Materials Discovery Project
OECD	Organisation for Economic Co-Operation and Development
OpenAIRE	Open Access Infrastructure for Research in Europe
RADAR	Research Data Repository
RatSWD	Rat für Sozial- und Wirtschaftsdaten
RDA	Research Data Alliance
RI	Research Infrastructure
SFB	Sonderforschungsbereich
SOEP	Das Sozio-oekonomische Panel
TIB	Technische Informationsbibliothek Hannover
VRE	Virtuelle Forschungsumgebung (Virtual Research Environment)
WR	Wissenschaftsrat

KURZFASSUNG

Der digitale Wandel hat Forschungsdaten und Forschungsmethoden erfasst, er verändert die Wissenschaft grundlegend. Der Umbruch verläuft unübersichtlich und hat eine dramatische Geschwindigkeit. Ihn gut zu steuern, stellt für Politik und Wissenschaft eine große Herausforderung dar. Der Prozess der Digitalisierung birgt Chancen. Er muss aber aktiv und auch durchgreifend gestaltet werden, um diese Chancen effizient zu nutzen. Ein leistungsfähiges Management von Forschungsdaten ist hierfür grundlegend wichtig, es ist der Schlüssel für Wissenschaft und für Innovation.

Mit seinem Positionspapier „Leistung aus Vielfalt“ spricht der Rat für Informationsinfrastrukturen (RfII) Empfehlungen aus, die zukunftsfähige Bedingungen für das Management von Forschungsdaten schaffen sollen. Gemäß dem durch die GWK erteilten Auftrag an den RfII betreffen diese Empfehlungen das Gesamtsystem von Wissenschaft in Deutschland. Ebenso sprechen sie verschiedene Akteure an – in der Politik wie in der Wissenschaft selbst. Aus Sicht des RfII wird es eine herausfordernde gemeinschaftliche Aufgabe sein, neue Formen des Umgangs mit Forschungsdaten sowie ihre Bewahrung und Nutzbarmachung für die Zukunft zu gestalten.

Das Positionspapier äußert sich zunächst zur Ausgangslage und zu existierenden wissenschaftspolitischen Spannungsfeldern. Aus Sicht des RfII ist der Umgang mit digitalen Forschungsdaten¹ in Deutschland trotz einiger guter Beispiele bislang noch durch überwiegend schwach koordinierte, projektförmige Initiativen geprägt. Eine Grundversorgung der Forschenden mit niedrigschwelligen Services für das Forschungsdatenmanagement fehlt, es bleibt derzeit bei Anstrengungen einzelner Institutionen und Organisationen. Kleinteiligkeit und befristete Finanzierung schaffen Effizienzprobleme. Mit hoher Personalfuktuation sind erhebliche Kompetenzverluste verbunden. Durch fehlende strategische Aufstellung werden Leistungsbereiche eingeschränkt, und im Hinblick auf wichtige Dimensionen wie Qualitätssicherung, Klärung von Rechtsfragen, Datenschutz und Datensicherheit drohen Nachteile im internationalen Wettbewerb. So besteht Handlungsbedarf in etlichen Feldern.

Der RfII leitet aus diesen Einschätzungen unterschiedlich geschnittene Empfehlungen ab:

¹ Der RfII verwendet einen weit gefassten Forschungsdatenbegriff, der sowohl analoge als auch digitale Sammlungen umfasst, vgl. hierzu die Begriffsklärung im Anhang A.

Fördermechanismen anpassen

Förderpolitisch empfiehlt der RfII, Projektförderungen von Forschungsdaten-Infrastrukturen nachhaltig auszurichten, da diese Art der Finanzierung für langfristig benötigte Dienste ein Risiko darstellt. Geschaffen werden muss ein transparenter Verstetigungsmechanismus. Ohne Fördergelder starr festzulegen, sollte Projektförderung so ausgestaltet werden, dass rechtzeitig Klarheit über Verstetigungsperspektiven und die hierzu erforderlichen Schritte besteht. Das Wertschöpfungspotenzial öffentlicher Investitionen wird auf diese Weise besser genutzt. Ziel von Verstetigungsprozessen muss es sein, ein System nachhaltiger Informationsinfrastrukturen zu schaffen, die sowohl verlässliche Arbeitsbedingungen für die Forschung bieten als auch eine dynamische Weiterentwicklung ermöglichen und eine qualitätsvolle Grundversorgung in Deutschland sichern.

Effizienz und Koordination durch eine (verteilte) nationale Infrastruktur

Als künftiges neues Rückgrat für das Forschungsdatenmanagement in Deutschland empfiehlt der RfII die Etablierung einer Nationalen Forschungsdateninfrastruktur (NFDI). Diese soll als bundesweites, verteiltes und wachsendes Netzwerk arbeitsteilig angelegt sein. Ein solches nationales Gefüge schrittweise aufzubauen, bietet die Chance, produktive Vielfalt durch Koordination und dringend notwendige Verknüpfungen in ein durchlässiges und weiterhin flexibles Gesamtsystem zu überführen. Die NFDI garantiert künftig Koordination, Kooperation und gemeinsame Standards. Als netzwerkförmige, dynamische Struktur verfügt die NFDI über Knoten unterschiedlicher Größe, die teils breit aufgestellte Servicezentren, teils spezialisierte Kompetenzzentren sein können. Dabei kann die institutionelle Anbindung variieren. Auch intelligente Lösungen für den schwierigen Bereich der Langzeitarchivierung/Langzeitverfügbarkeit von Forschungsdaten gehören zu den Aufgaben der NFDI.

Datenverantwortung, Datenkultur

Die Umstellung auf digitale Prozesse bedeutet, dass sich auch fast alle ‚weichen‘ Rahmenbedingungen für Wissenschaft ändern – und das betrifft Verantwortlichkeiten im Handeln der Forschenden selbst. Für die Bereiche der Qualitätssicherung digitaler Forschungsdaten, für die rechtlichen Rahmenbedingungen der Datennutzung (Stichwort ‚Open Science‘) sowie für die Felder Datenschutz und Datensicherheit spricht der RfII eine Reihe von Empfehlungen aus. Durch diese soll Datenverantwortung in der Wissenschaft für alle Phasen des ‚Datenlebenszyklus‘ klar verankert werden. Politik und Wissenschaft müssen verstehen und konkret machen, was gute wissenschaftliche Praxis im Digitalzeitalter heißt. Gefordert sind hier auch die Wissenschaftsorganisationen: Flexible Organisationsformen, Monitoring und Evaluationen gilt es im Wissenschaftssystem so auszugestalten, dass sie Anreize für gutes Forschungsdatenmanagement setzen und auf Seiten von Forschenden wie auch öffentlich Vertrauen schaffen. Wissenschaftliche Leistungsfähigkeit, gute Datenkultur und gesellschaftliche Wertschöpfung gehören zusammen.

Monitoring und Qualitätssicherung

Wichtig ist ebenso die Empfehlung, auf allen Ebenen des Wissenschaftssystems die Personalentwicklung voranzutreiben. Datenintensive Forschung wie auch Lehre benötigen einschlägig qualifizierte Personen – und dies in großer Zahl. Hier läuft der Arbeitskräftemarkt in Deutschland der globalen Entwicklung in gefährvoller Weise hinterher. Daher gilt es aus Sicht des RfII, schnellstmöglich kompetente Forschergenerationen und spezialisierte Fachkräfte für neue Berufsbilder im Bereich des Datenmanagements auszubilden. In gleicher Weise müssen Bewusstsein und Handlungskompetenzen auf den Managementebenen befördert werden: Zur Lösung von Infrastrukturproblemen sind vielfach Kommunikations- und Prozessmanagement entscheidend. Geeignete Kompetenzbeschreibungen, attraktive neue Berufe und Studiengänge müssen daher dringend geschaffen werden. Es sind die ‚Köpfe‘, die durch Integration von Daten, Informationen und Wissen aus heterogenen Quellen und über Domänengrenzen hinweg für die Entstehung neuen wissenschaftlichen Wissens und damit für die Wertschöpfung aus den Daten sorgen.

Personalentwicklung
auf allen Ebenen

Ein besonderes Gewicht hat für den RfII über alle Empfehlungen hinweg die internationale Dimension. Datenflüsse machen (wie Forschung) an Grenzen nicht Halt. Forschungsdatenmanagement in Deutschland muss sich im europäischen und globalen Rahmen entwickeln. Dies meint nicht allein Wettbewerb, sondern Lernen, Kooperation und aktive Mitgestaltung. Der RfII sieht Deutschland hierbei in einer wichtigen Rolle. Er ermutigt Politik und Wissenschaft ausdrücklich zu einer noch stärkeren internationalen Vernetzung.

Verstärkte
internationale
Vernetzung

Umsteuern ist mit Aufwand verbunden. Der RfII geht davon aus, dass zur Sicherung der Leistungsfähigkeit des deutschen Wissenschaftssystems tatsächlich erheblicher Investitions- und Finanzierungsbedarf auf verschiedenen Ebenen des Gesamtsystems besteht. Die anstehenden Aufgaben betreffen die Gestaltung eines mehrjährigen Übergangsprozesses: Wissenschaft muss sich in fast allen Bereichen ändern, wenn weltweit aus Forschungsdaten digitale Forschungsdaten und aus Forschungsmethoden digitale Forschungsmethoden werden. Dies gilt auch für Infrastruktureinrichtungen, Repositorien, Archive etc., deren Arbeitsschwerpunkte sich anpassen und entwickeln müssen. Der RfII appelliert in dieser Schwellensituation an die Entscheidungsträger in Wissenschaft und Wissenschaftspolitik, kraftvolle Sonderanstrengungen zu unternehmen, um die Weichen zeitnah richtig zu stellen. Gemeinsam mit den Ländern ist deshalb auch der Bund gefordert, zur wirksamen und raschen Überwindung der derzeit vorherrschenden Zerfaserung der die Forschungsdaten betreffenden Aktivitäten aktiv beizutragen. Ohne Mobilisierung weiterer Ressourcen in der Wissenschaft selber und zusätzlich durch die Zuwendungsgeber wird dieses Ziel nicht zu erreichen sein.

Übergangsprozess
gestalten

Der Rfll betont, dass die Herausforderungen komplex sind, weswegen nur ein Zusammenwirken zum Erfolg führt. Akteure in der Wissenschaftspolitik und Akteure auf Seiten der Selbstorganisation der Wissenschaft tragen jeweils in ihren Bereichen Verantwortung dafür, dass es zu gemeinsamem Handeln kommt. Für alle muss es entscheidend sein, den Wandel zu gestalten, Deutschland in einer Situation des globalen wissenschaftlichen Wettbewerbs gut zu positionieren, öffentliche Mittel möglichst effizient einzusetzen und ein maximales Wertschöpfungspotenzial für Wissenschaft, Wirtschaft und Gesellschaft nachhaltig zu realisieren.

EXECUTIVE SUMMARY

The digital revolution is fundamentally transforming research data and methods; indeed, science on the whole is in a state of flux. Mastering this transformation poses major challenges for actors in the domains of science and policy, in part because the precise nature of the transformation is not yet clear. The process of digitalization creates immense opportunities, but it must be structured proactively. To this end, the establishment of effective governance mechanisms for research data management (RDM) is of fundamental importance and will be one key driver for successful research and innovation in the future.

In its position paper, the German Council for Scientific Information Infrastructures (RfII) makes a series of recommendations concerning how research data should be managed in the future. The RfII was tasked by Germany's Joint Science Conference (GWK) with formulating broad-based recommendations for the science system in Germany as a whole. Consequently, the recommendations presented here have ramifications for a variety of actors in the domains of politics and science. The RfII is convinced that the adoption of new ways in dealing with research data as well as long-term preservation and accessibility will be a significant, common challenge for all actors in the years ahead.

The position paper describes current policies and practices for managing research data and discusses a number of conflicting priorities in science policy. While there are several good examples of research data management¹ in Germany, there is an overall absence of coordination, and current efforts often take the form of parallel, project-based initiatives. Universal access to services for data management is lacking, as the key actors at present are individual institutions and organisations, and their efforts often suffer from limited financing and/or excessive niche focus. High staff turnover means that valuable know-how is frequently lost. Furthermore, the range of services being provided is impaired by the absence of governance mechanisms which could impart greater strategic direction. In addition, there is a risk of international competitive disadvantage for Germany due to unresolved issues in the domains of quality assurance, legal compliance, data privacy, and data security. In this way, there is a clear need for action in a variety of areas.

Based on the foregoing findings, the RfII has developed a series of recommendations:

¹ The RfII has a broad conception of research data that encompasses both analogue and digital collections.

Adjusting funding mechanisms

With regard to funding policy, the RfII recommends implementing long-term funding mechanisms in line with the long-term nature of research data management. Current project funding schemes do not foster long-term sustainability, thus jeopardizing the development of services needed in the long run. Policy-makers need to establish clear pathways for institutions and organizations to obtain the resources they require over the long term. Without tying up subsidies in a fixed manner, funding phases for RDM infrastructures should be designed so that applicants have a clear understanding of their chances for securing long-term financing and the steps that are necessary to obtain it. This will boost the value that is derived from public investment. The goal of funding policy should be to develop an ecosystem of sustainable infrastructures that provides researchers in Germany with universal and reliable access to data management services. Such an ecosystem should also be designed to grow dynamically over time.

Efficiency and coordination based on a (distributed) national infrastructure

The RfII recommends the establishment of a *Nationale Forschungsdateninfrastruktur* (National Research Data Infrastructure, or NFDI), which will serve as the backbone for research data management in Germany. The NFDI should be implemented as a national collaborative network that grows over time and is composed of various specialized nodes. The establishment of such a network is recommended on a step-by-step basis, as this will ensure the overall management system remains flexible while also facilitating the productive integration of diverse resources. The NFDI will provide for coordination, cooperation, and common standards. As a network-based, dynamic organizational structure, the NFDI will be composed of nodes of various sizes. Some of these nodes will take the form of broad-based “service centres”, while others will be specialized “centres of excellence” for specific subareas. These centres can be established within various existing organisations and institutions. The difficult issue of developing intelligent solutions for the long-term archiving/provisioning of research data will be among the areas addressed by the NFDI.

Responsible data culture

The current transition to digital processes means that nearly all of the “soft” factors for scientific activity are in a state of flux. The responsibilities that fall to researchers have to be recalibrated. The RfII thus makes various recommendations regarding data quality assurance, the adoption of a legal framework for data reuse (based on the Open Science model), and data privacy and protection. These recommendations aim to define the responsibilities borne by researchers and their organisations during all phases of the “data life cycle”. Accordingly, policy-makers and scientists need to understand and set forth good scientific practice for research in the digital age. Scientific organisations also have a clear role to play in this area: monitoring and evaluation systems should be designed to create incentives for good research data management while also

Monitoring and quality assurance

engendering trust among researchers and the lay public. Good data management practices go hand in hand with research that is cutting edge and has a strong practical value for society as a whole.

The Rfll additionally recommends that due attention be devoted at all levels to human resources development. Adequately qualified individuals are required in large numbers for data-intensive research and teaching. The German labour market is dangerously lagging behind global developments in this area. Accordingly, the Rfll sees a pressing need to educate a new generation of highly capable researchers and specialist employees for new occupations in the area of data management. At the same time, new skills and greater awareness need to be promoted at management levels, as communication and process management are frequently decisive for solving infrastructure problems. Consequently, there is an urgent need to develop new occupational profiles and fields of study. Ultimately, it is the individuals behind the system, with their skills and aptitudes, who generate new scientific knowledge and added value by integrating data, information, and knowledge from heterogeneous sources.

Comprehensive human resources development

The Rfll attaches particular importance to international collaboration, for the flow of digital information (like research) knows no borders. The RDM ecosystem in Germany will have to develop within a broader European and global context. This does not only mean competition, but more importantly active collaboration and mutual learning. The Rfll believes that Germany has a particular role to play in this area, and calls upon policy-makers and scientists to develop closer networking ties to their counterparts in other nations.

Stronger international ties

Changing course requires time, cost, and effort. The Rfll believes that considerable investments are required at various levels of the German scientific system in order to ensure its future efficacy. The tasks in need of attention will require a long transition process. The scientific system will need to change in nearly all areas if research data and methods are to become truly digital in coming years. The need for change also applies to existing infrastructures, repositories, and archives, which need to reorient towards new tasks. At this current tipping point, the Rfll calls upon decision-makers in the domains of science and policy to make vigorous efforts to facilitate the necessary transition phase. This call for action extends to Germany's *Länder* and the national government, which are urged to undertake effective and rapid measures to overcome the current fragmentation of efforts in the area of research data management. This fragmentation cannot be overcome without mobilizing substantial resources within the science system, including its funders.

Actively steering the transition

The Rfll emphasizes that the challenges are complex and that collaboration will be decisive for success. Scientific policy-makers and the scientific community are equally responsible for ensuring that joint action takes place.

For all those involved, the overarching concern should be to support the international competitiveness of German scientific research and make efficient use of public funding. With effective collaboration Germany can actively push the digital transformation to ensure maximum benefit for science, industry, and society.

1 HERAUSFORDERUNGEN UND AUFGABENSTELLUNG

Mit dramatischer Geschwindigkeit setzt sich Digitalität der Datenerhebung, der Arbeitsverfahren sowie der Publikations- und Verwertungswege in der Wissenschaft durch. Sinnfällig stand in diesem Zusammenhang zunächst die Großgeräteforschung im Fokus von Wissenschaftspolitik und Forschungsfinanzierung. Herausforderungen wie die Digitalisierung analoger Wissensbestände und den Zugang zu digitalem Wissen betreffende Agenda-Punkte wie *Open Access* oder *Informationskompetenz* kamen hinzu.

Deutlich grundsätzlicher rücken inzwischen für das Gesamtfeld von Forschung, Lehre und Transfer Fragen nach Forschungsdaten und nach dem zukünftigen Management digitaler Forschungsdaten nach vorn. Schon aufgrund ihrer enormen Menge und Heterogenität – bei zugleich hoher Spezifik und vergleichsweise raschem Verfall – herrscht Handlungsbedarf. Sowohl eine optimale Erschließung als auch die Bewahrung und Nutzbarmachung dieser Daten bedarf politisch gesetzter Rahmenbedingungen, die es derzeit nicht gibt. Die wissenschaftliche Nachnutzung qualitätsgesicherter Forschungsdaten schafft einen erheblichen Mehrwert öffentlicher Investitionen in die Forschung und birgt ein hohes Wertschöpfungspotenzial für Wissenschaft und Wirtschaft. Der digitale Wandel ist durch systemrelevante Schwellen gekennzeichnet, an denen zur Ermöglichung leistungsfähiger neuer Strukturen Pfadentscheidungen gefragt sind. Unterlassungen haben an solchen Schwellen negative Folgen. Man bleibt nicht bloß zurück, sondern Fehlentwicklungen drohen.

Der Ruf nach *Forschungsdatenmanagement* (FDM) beinhaltet allerdings ein komplexes Geflecht von Erfordernissen und Gestaltungsfragen. Forschungsdaten sind nicht allein die (End-)Ergebnisse von Forschung. Es handelt sich vielmehr um jegliche Daten, die im Zuge wissenschaftlichen Arbeitens entstehen, einschließlich zahlreicher Daten, die z. B. im Zuge von Messung, Selektion, Aufbereitung, Sammlung und Aufbewahrung einer möglichen Belegfunktion für wissenschaftliche Vorhaben dienen. Das Management *digitaler* Forschungsdaten umfasst ein breites Spektrum von – über das Forscherhandeln im engeren Sinne hinaus auch organisationsbezogenen – Maßnahmen. Diese müssen getroffen werden, um Ergebnisse auch in der digitalen Welt reproduzierbar und Daten zur Nachnutzung verfügbar zu machen, und zwar im gesamten *Datenlebenszyklus* über Erhebung, Verarbeitung, Analyse und Archivierung hinweg.²

Seit der Jahrtausendwende sind international zahlreiche Initiativen zum Aufbau von Datensammlungen, zum Hosting, zu Standardisierung, Verrechtlichung und

² Der Rfll hat zu diesen und weiteren häufig verwendeten Ausdrücken Begriffsklärungen erarbeitet. Sie sind dem Positionspapier als Anhang A beigefügt.

Verwertung zu beobachten. Auch in Deutschland gibt es hierzu Aktivitäten, und es liegt eine Vielzahl von Empfehlungen, Agenden und Appellen von Seiten wichtiger Akteure des deutschen Wissenschaftssystems vor.

Die gegenwärtige Situation ist im Positiven von *Vielfalt*, im Negativen von *Zersplitterung* gekennzeichnet. Der Wandel kennt viele ungleichzeitige Entwicklungen und unterschiedliche Entwicklungsstände. Forschende und Servicepartner, Hochschulen und außeruniversitäre Einrichtungen, Disziplinen und Projektverbände, Forschungsförderer und Ministerien experimentieren mit Optionen, die sich meist *bottom up* ergeben. Entscheidungsträger in Wissenschaft und Politik haben lange Zeit den Aufbau neuer Strukturen *bottom up* stärker begünstigt als deren Konsolidierung und übergreifende Systematisierung. Viele der selbstorganisierten Infrastrukturen sind noch im Experimentierstadium und technisch wie organisatorisch wenig kompatibel. Der Erhalt aller dieser Systeme ist auf Dauer nicht finanzierbar; ihrer nachträglichen Zusammenführung sind technisch, organisatorisch und finanziell Grenzen gesetzt; sich abzeichnende Aufwände müssen in den kommenden Jahren durch Interventionen sinnvoll reduziert werden. Eine Konsolidierungsphase sollte daher vorausschauend geplant werden, um die Spitzenstellung der deutschen Wissenschaft zu sichern.

Der rasch voranschreitende internationale Wettbewerb ist mehr als eine Randbedingung für den Umgang mit Forschungsdaten. Forschung in Deutschland ist in hohem Maße europäisch und global vernetzt, sie benötigt breiten Zugang zu weltweiten Ressourcen, um wettbewerbsfähig zu sein. Aber auch private Angebote von Software, Kommunikations-, Speicher- und Publikationsdiensten werden durch Forschende in Deutschland massenhaft genutzt. In neuem Ausmaß gelangen dabei auch Forschungsdaten ins Internet und werden in Zugriffssphären transferiert, für die deutsches Recht nicht gilt. Wissenschaftspolitisch geben auch restriktive Publikations-/Abonnementsmodelle kommerzieller Anbieter Anlass zur Sorge. Das Open-Science-Paradigma einer weltweiten Verfügbarkeit von Datenbeständen hat ebenfalls eine Kehrseite. Welche Auswirkungen maximale *Offenheit* von Forschungsdaten auf die wissenschaftliche und wirtschaftliche Wettbewerbsfähigkeit der jeweiligen Datenproduzenten in Nationalstaaten, Regionen und Institutionen haben wird, ist bislang wenig diskutiert.

An Hinweisen auf fehlende Regelungen, Ressourcen und Infrastrukturen wie auch an Empfehlungen für künftige Entwicklungen im Forschungsdatenmanagement fehlt es nicht. Dennoch fällt es Entscheidern in Wissenschaft und Politik schwer, Ansatzpunkte für Maßnahmen zu finden sowie in einer strategischen Perspektive in Dienste und Forschungsdaten-Infrastrukturen zu investieren. Dies ist nachvollziehbar, denn Anforderungen an das Management digitaler Forschungsdaten betreffen tatsächlich fast alle Kernprozesse von Forschung, Lehre und Wissenstransfer. Sie laufen auf einen Umbau bestehender Organisation hinaus. Auf die Herausforderungen umfassend zu reagieren, wird

Wissenschaft also in mehr oder weniger starker Weise verändern. Nicht nur deshalb sind Steuerungsentscheidungen mit Verantwortung verbunden.

Für die aus guten Gründen dezentrale, netzwerkartige Struktur des deutschen Wissenschaftssystems mit seiner Vielfalt an Disziplinen, Institutionen und Forschungsformen ist eine bedarfsgerechte Dateninfrastruktur nicht *top down* planbar. Allerdings sind entschlossene Top-down-Impulse gefragt. Der seitens des Wissenschaftsrates (WR) und der Gemeinsamen Wissenschaftskonferenz (GWK) gleichermaßen favorisierte Grundsatz eines Vorrangs der „stimulierten Selbstorganisation“³ muss auch für das komplexe Feld des Umgangs mit Forschungsdaten gelten. Gleichwohl bedarf es hierfür eines Rahmens. Somit zielen die durch den Rfll hiermit vorgelegten Empfehlungen auf ein besser *orientiertes* sowie stärker *koordiniertes* und auch nachhaltig *unterstütztes* Verhalten der verschiedenen Akteure im Sinne eines lernenden Systems. Es gilt, die Balance in folgenden Spannungsfeldern zu wahren:

- Projektförderung vs. Institutionalisierung
- *Top down* vs. *bottom up* koordinierte Prozesse
- Open Data vs. Datensouveränität zur Eigenforschung (Kooperation vs. Konkurrenz)
- Disziplinäre bzw. institutionenspezifische Lösungen vs. übergreifende Dienste
- Datenschutz vs. Forschungsfreiheit
- Staatliche Förderung vs. privatwirtschaftliches Engagement
- Infrastrukturleistungen behördlicher Art vs. Angebote von Großforschungs- oder Hochschuleinrichtungen
- Investitionen in Infrastrukturen vs. Investitionen in Köpfe
- nationale Aktivitäten vs. europäisches und globales Umfeld

Dringend und also kurzfristig zu überwinden sind praktische Hürden, die in Förderbedingungen, institutionellen Barrieren und mangelnder Anerkennung von Infrastrukturleistungen liegen. Ebenso sind rechtliche Problemstellungen in einer Weise zu klären, die den Bedarfen der Wissenschaft entspricht. Studien zeigen zudem, dass viele Forschende die neuen Möglichkeiten noch kaum oder gar nicht nutzen. Angesichts dieser praktischen und kulturellen Hürden bedarf es klarer Regeln, eines Zielsystems und geeigneter Unterstützungsfunktionen. Die neue Struktur sollte eine indirekte Koordination ermöglichen und fördern sowie Kontrollpunkte und Mechanismen etablieren, die *ex post* zusammenführen, was zusammen gehört.

³ GWK (2013) – Drucksache 13.48; WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 81.

2 AUSGANGSLAGE

2.1 FORSCHUNGSDATEN, INFORMATIONSINFRASTRUKTUREN UND WISSENSCHAFTSPOLITIK

Schlüsselstellung der Infrastrukturen

Datenmanagement benötigt Infrastruktur – an diesem Punkt ist die Wissenschaftspolitik zur Gestaltung des Wandels gefordert. Forschungsinfrastrukturen waren für die Wissenschaftsentwicklung seit je her von zentraler Bedeutung. Seit im Zuge wissenschaftlicher Vorhaben zunehmend digitale Forschungsdaten entstehen und diese insbesondere (z. B. durch Messung, Simulation, Selektion, Aufbereitung, Sammlung und Aufbewahrung) einer möglichen Belegfunktion dienen, rückt die besondere Bedeutung von Informationsinfrastrukturen nach vorn. Der Begriff bezeichnet technisch und organisatorisch vernetzte Dienste und Angebote für die Arbeit mit Daten-, Informations- und Wissensbeständen in der Wissenschaft.⁴ Diese sind in der Regel eng mit digitalen Methoden, Prozessen, fachspezifischen Services und Forschungsformen verbunden. Die Rede ist also nicht nur von Geräten und Datenbanken, sondern vom Forschungsprozess selbst. Denn fast alle Forschungswerkzeuge haben heute relevante IT-Komponenten, die Forscherkommunikation erfolgt digital vernetzt und auch die wissenschaftliche Informationsversorgung durch Bibliotheken, Informationszentren und Archive hat sich auf Digitalität umgestellt.

Sowohl für die Bewältigung der schiereren Menge als auch für den digital veränderten Umgang mit (und den nutzenden Zugang zu) digitalen Forschungsdaten sind Informationsinfrastrukturen – im geschilderten weiten Sinne von Diensten und Verfahren – ein Schlüssel. Rechtliche und die Wissenschaftskultur betreffende Aspekte kommen hinzu.

Akteure des Wissenschaftssystems haben daher in den letzten Jahren mit einer Vielzahl an Stellungnahmen, Studien und Positionspapieren zu grundlegenden Themen auf den digitalen Wandel reagiert. Solche Papiere behandelten u. a. das Open-Data-Paradigma, den systematischen Aufbau von Dateninfrastrukturen, geforderte Anreize für einen Kulturwandel in der Wissenschaft sowie neue Berufsbilder (vgl. auch Anhang B).⁵

⁴ Informationsinfrastrukturen, im Englischen (etwas enger gefasst) „e-Infrastructures“, sind als eine spezifische Form von Forschungsinfrastrukturen zu verstehen. Vgl. die Begriffsklärungen im Anhang A.

⁵ Bericht des Redaktionsausschusses Konzepte: „Die Entwicklung von Konzepten für Informationsinfrastrukturen in der Bundesrepublik Deutschland seit den 1960er Jahren“.

Eine High Level Expert Group der Europäischen Kommission postulierte 2010 eine kollaborative Dateninfrastruktur („Collaborative Data Infrastructure“), die einen übergangslosen Zugriff auf vertrauenswürdige Daten sowie deren Nutzung über verschiedene Systeme hinweg unterstützt.⁶ In der Folge entstanden auch mit deutscher Beteiligung EU-geförderte Initiativen, die auf eine solche Infrastrukturentwicklung hinarbeiten, wie die Research Data Alliance (RDA), European Data Infrastructure (EUDAT) und die geplante European Open Science Cloud (EOSC).⁷

In Deutschland intensivierte sich der Diskurs nach Erscheinen des Gesamtkonzepts der *Kommission Zukunft der Informationsinfrastruktur* (KII) und einer Serie verbundener Empfehlungen des Wissenschaftsrats in den Jahren 2011 und 2012.⁸ Die Deutsche Initiative für Netzwerkinformationen (DINI) als Verband der Bibliotheken, Rechen- und Medienzentren, die Deutsche Forschungsgemeinschaft und die Hochschulrektorenkonferenz publizierten Stellungnahmen, die auch die jeweiligen institutionellen Rollen reflektieren. Aus der Perspektive der Forschung stehen Fragen der Zugänglichkeit und der Sicherung absehbar wachsender Datenmengen im Vordergrund. Datenmanagement und datenbasierte Wissenschaft – einschließlich digitaler Analysemethoden – müssen unterstützt und erleichtert werden. Zentrale, wiederholt benannte Herausforderungen sind die nachhaltige Finanzierung geeigneter Infrastrukturen, die Aus-, Fort- und Weiterbildung des Fachpersonals in Wissenschaft und Informations-einrichtungen sowie die Entwicklung begleitender Regelwerke, Policies und Selbstverpflichtungen, aber auch die Konkretisierung gesetzlicher Bestimmungen.⁹

Das Positionspapier der Arbeitsgruppe Forschungsdaten in der Schwerpunkti-niative *Digitale Information* formuliert für die Wissenschaft eine Vision im Geiste der *Berliner Erklärung*: Bis 2025 sollen „Wissenschaftlerinnen und Wissenschaftler aller Disziplinen [...] auf alle Forschungsdaten einfach, schnell und ohne großen

Positionspapiere
in Deutschland
und der EU

⁶ European Commission (2010) – Riding the wave, S. 4: „Our vision is a scientific e-infrastructure that supports seamless access, use, re-use, and trust of data. In a sense, the physical and technical infrastructure becomes invisible and the data themselves become the infrastructure [...]“.

⁷ Research Data Alliance – <https://rd-alliance.org>; EUDAT – <https://www.eudat.eu/what-eudat>; EOSC – <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud> (zuletzt geprüft am 25.04.2016). Zur European Cloud Initiative vgl. auch European Commission (2016) – Data and knowledge economy und BMBF (2016) – Eckpunkte European Cloud Initiative (unveröffentlicht).

⁸ Vgl. KII – Kommission Zukunft der Informationsinfrastruktur (2011) – Gesamtkonzept und WR (2012) – Empfehlungen zu Informationsinfrastrukturen (sowie die darin referenzierten Empfehlungen der Vorjahre).

⁹ Zur Entwicklung des Diskurses siehe ausführlicher Anhang B, dort auch vollständiger Literaturnachweis.

Aufwand zugreifen [können], um auf höchstem Niveau zu forschen und exzellente Ergebnisse zu erzielen.¹⁰

Die Bundesregierung hat in der 2014 verabschiedeten Digitalen Agenda ebenfalls den besseren Zugang zu Forschungsdaten als Ziel verankert. Die Agenda wie auch weitere Maßnahmen werden dabei als Teil einer Internationalisierungsstrategie für die Wissenschaft verstanden.¹¹

Sie erkennt damit Handlungsbedarf für ein Feld, das in der digitalen Frühphase der Bundesrepublik in den 1970er Jahren zunächst über sogenannte „Fachinformationssysteme“ für die Wissenschaft zentralstaatlich organisiert war, seit den 1980er Jahren dann aber stärker in die Verantwortung der Wissenschaft selbst überging. Die mit Liberalisierung und Eigenverantwortung verbundenen Chancen konnten die wissenschaftlichen Akteure im digitalen Wandel aktiv nutzen. Verteilte Aktivitäten und wettbewerbliche Elemente haben für eine dynamische Entwicklung gesorgt.

Vielfalt projektförmi- ger Lösungsansätze in Deutschland

Inzwischen hat sich – stimuliert auch durch Förderprogramme – eine weitgehend dezentrale ‚Landschaft‘ von Lösungen und auch Infrastrukturen herausgebildet (vgl. 2.2). Diese sind teils auf datenbezogene Services (Speichern, Rechnen, Transferieren), teils mehr auf informationelle Dienstleistungen (erschließendes Archivieren, Informieren, Recherchieren), teils auf bestimmte wissenschaftliche Methodiken bzw. Forschungsformen (Beobachten, Experimentieren, Messen, Simulieren, hermeneutisches Interpretieren, theoretische Analyse, Design/Gestalten) ausgerichtet.¹²

Der Status quo der auf diese Weise stimulierten Selbstorganisation in der Wissenschaft wirft allerdings Probleme auf. Entwicklungen verlaufen häufig ungleichzeitig und inkrementell, heterogene Akteure sind oft unabgestimmt tätig, die Herausbildung übergreifender qualitätssichernder Standards¹³ ist mühsam, Investitionen lassen sich kaum koordinieren und Ressourcenprobleme sorgen für Fehlentwicklungen bzw. Stagnation. Vielfalt bleibt ungesteuert, Synergien durch eine übergreifende Daten(nach)nutzung lassen sich nicht heben. In Wissenschaft und Politik besteht deswegen gleichermaßen ein hohes Interesse an Projekten zur Konsolidierung und Systematisierung. Viele Dienste erscheinen zudem durch den vorherrschenden Modus der Projektfinanzierung für die Forschenden noch wenig verlässlich. So bleibt teils die Nutzung von Diensten und Ressourcen wie auch die Akzeptanz des Forschungsdatenmanagements in der Breite noch hinter den

¹⁰ Allianz-Initiative Digitale Information – AG Forschungsdaten (2015) – Research data at your fingertips, S. 1.

¹¹ Wissenschaftspolitische Ziele der Digitalen Agenda finden sich wortgleich in der Strategie der Bundesregierung zur Realisierung des Europäischen Forschungsraums. Vgl. BMBF (2014) – Strategie EFR; Bundesregierung (2014) – Digitale Agenda 2014-2017.

¹² Vgl. WR (2012) – Empfehlungen zu Informationsinfrastrukturen, S. 35.

¹³ Zum Verständnis von Standards vgl. auch die Begriffsklärung im Anhang A.

Erwartungen zurück – was dem Forschungsstandort Deutschland im Wettbewerb schadet.

Aus Sicht des Rfll liegen zahlreiche informative Stellungnahmen und Problemdiagnosen rund um das Politikfeld eines zukunftsfähigen Forschungsdatenmanagements bereits vor.¹⁴ Anzugehen sind somit vor allem Punkte, die das angesichts der Komplexität der Lage zu beobachtende Vollzugsdefizit trotz existierender Problembeschreibungen beheben.

Problembeschreibungen vs. Vollzugsdefizit

Dabei ist – neben der Komplexität der Anforderungen – auch die gegebene institutionelle Ausgangslage in Betracht zu ziehen. Sowohl das Bund/Länder-Gefüge als auch die Balance zwischen Universitäten/Hochschulen und außeruniversitären Forschungseinrichtungen einschließlich der Ressortforschung sollten bei anstehenden wissenschaftspolitischen Entscheidungen mitgedacht werden. Ebenso ist auf der Strukturebene die überkommene Arbeitsteilung von Rechenzentren, Bibliotheken und eher „thematischen“, disziplinen- oder methodengebundenen Datenzentren bzw. Archiven in den Blick zu nehmen. Die Arbeitsprofile von Informationsinfrastrukturen sollten bestmöglich aufeinander abgestimmt sein.

2.2 ÜBERBLICK ÜBER DIE „LANDSCHAFT“

Digitale Forschungsdaten werden in der vollen Breite des Wissenschaftssystems generiert. Bestimmte Formen datenintensiver Großgeräteforschung (z. B. Kernphysik, Erdbeobachtung, Astrophysik, Klimamodellierung) sind typischerweise in der *außeruniversitären Forschung* anzutreffen. Aber auch die Forschung an *Hochschulen* produziert in enormen Umfängen digitale Daten (z. B. Genomik/Proteomik, Pharmaforschung mit sog. Hochdurchsatz-Screening, Archäologie/Altewissenschaften, Forschung mittels Bildgebung, Audio- und Videodaten, Sprachkorpora). Auch sind die Hochschulen an der Erstellung von großen Datennetzwerken beteiligt (z. B. Geodaten, medizinische Versorgungsdaten), die Fächer- und Projektvielfalt produziert zudem viele kleinteilige und sehr heterogene Datenbestände.¹⁵ Hinzu kommen die *wissenschaftlichen Bibliotheken, Sammlungen, Archive* und *Museen* (z. B. Bohrproben, Biomaterialien, Kunstwerke), die insbesondere durch die rückwärtige Digitalisierung von Beständen (z. B. Texte, Bilder, Kulturgüter, naturkundliche Archivalien) große Datenmengen erzeugen. Schließlich sind Daten aus der *Ressortforschung* und von Messungen und Erhebungen *behördlicher Stellen* zu nennen; diese fallen in großem Umfang auf den unterschiedlichen Ebenen der Verwaltung und in vielen nichtstaatlichen Einrichtungen an (z. B. Umweltdaten, Landwirtschaftsdaten, Sozialdaten, Wirtschaftsdaten, Krankenkassen-

Wo entstehen Forschungsdaten?

¹⁴ Vgl. Rfll (2015) – Auftakterklärung, S. 3.

¹⁵ Eine Übersicht der damit verbundenen Herausforderungen gibt Anhang C (Szenarien des Forschungsdatenmanagements).

Arbeit mit Forschungsdaten

daten, Unternehmensdaten, Daten aus Telekommunikationsanwendungen etc.). All diese Daten sind in variablen Kombinationen für die Forschung unentbehrlich.

Ähnlich komplex stellt sich die Lage in der Arbeit mit Forschungsdaten dar. Hier ist insbesondere anzumerken, dass Digitalität der standortübergreifenden Verbundforschung (EU-Konsortien, Sonderforschungsbereiche, Exzellenz-Cluster, Graduiertenschulen etc.) entgegenkommt. Namentlich in interdisziplinären Konstellationen treffen dabei *heterogene Methoden und Forschungsformen* mit ihren jeweils verschiedenen Datentypen und Anforderungen an die Datenqualität sowie die Datenprozessierung aufeinander. In den letzten Jahren wurden zudem verstärkt sogenannte *Virtuelle Forschungsumgebungen* (VRE) etabliert; diese haben das Ziel, digitale (Einzel-) Werkzeuge maßgeschneidert zu integrieren und Ressourcen zu vernetzen. Zugleich sollen sie *Kommunikation zwischen Forschenden* eng am Forschungsprozess entlang unterstützen. Als niedrigschwellige Dienste sind in diesem Zusammenhang freie Web-2.0-Werkzeuge zu nennen (Lab Archives, Github oder auch Cloud-Dienste wie Dropbox), die disziplinenübergreifend für den Austausch von Daten genutzt werden. Hoch integrierte digitale Forschungsumgebungen setzen typischerweise eher homogene Methoden und Communities¹⁶ voraus. An Hochschulen kommen Forschungsdaten überdies auch in der *wissenschaftlichen Lehre* zum Einsatz; die Daten sollten also in digitale Lernumgebungen integrierbar sein.

Digitale Publikation von Forschungsdaten

Für die digitale Publikation von Forschungsergebnissen gilt, dass Beiträge – zumal in empirisch oder experimentell arbeitenden Disziplinen – zunehmend mit Forschungsdaten verknüpft werden. Auch die Publikation von Forschungsdaten beginnt sich zu etablieren, z. B. in sogenannten „Data Journals“ in den Erd- und Umweltwissenschaften, der Nanotechnologie oder der sozialwissenschaftlichen Forschung.¹⁷ Institutionell spielen hier einerseits klassische *Verlage* eine Rolle, wobei wenige globale Akteure das Wissenschaftsverlagswesen ökonomisch dominieren. Andererseits werden internationale Fachzeitschriften auch von *Fachverbänden* (so z. B. in der Physik) publiziert und auch die Bedeutung der digitalen *Selbstpublikation* steigt. Im Umfeld von Verlagen entstehen darüber hinaus Dienstleister wie das durch Mitgliedsorganisationen finanzierte Datenrepositorium Dryad oder das Unternehmen Figshare: Diese bieten Autorinnen und Autoren die – zumeist gebührenpflichtige – Publikation der zu ihren Aufsätzen gehörenden Daten an.

¹⁶ Zum Community-Begriff vgl. Anhang A.

¹⁷ Vgl. Beispiele von Data Journals unter <http://proj.badc.rl.ac.uk/preparde/blog/DataJournalsList> (zuletzt geprüft am 21.04.2016), in den Erd- und Umweltwissenschaften Pangaea: <https://www.pangaea.de> (zuletzt geprüft am 25.04.2016).

Ebenfalls zur Selbstpublikation genutzt werden wissenschaftliche *Online-Netzwerke* (Academia.edu, Mendeley, ResearchGate), die teils unter verlagsartigen Bedingungen variable Dienstleistungen für registrierte Mitglieder offerieren.¹⁸

Anbieter von Diensten für den Austausch, die informations- und nutzungsorientierte Sammlung, Aufbewahrung, Erschließung, Aufbereitung und ggf. auch Bewirtschaftung wissenschaftlicher Daten und Informationen sind typischerweise *Bibliotheken* aber auch Archive sowie thematisch oder fachlich ausgerichtete *Forschungsdatenzentren*, *Forschungsinformationszentren* und *Repositorien*. Die deutsche Landschaft umfasst hier eine Fülle von Einrichtungen mit unterschiedlichen Trägerschaften, Konsortialstrukturen und kommerziellen Vertragspartnern, Dienste-Portfolios, Verwertungskonzepten sowie Nutzergruppen aus Wissenschaft und Wirtschaft. Einige deutsche Datenzentren oder-sammlungen sind Bestandteil verteilter transnationaler Infrastrukturen, die sich aus verschiedenen nationalen „Knoten“ speisen (etwa GBIF, NoMaD, ELIXIR, im Kulturbereich das Repositorium Europeana). Wo sich solche „Knoten“ formiert haben, hat dies nicht selten eine Konsolidierung der entsprechenden Einzelaktivitäten auf nationaler Ebene bewirkt (siehe 2.4). Auf der europäischen Ebene ging 2014 die ebenfalls als verteiltes System konzipierte wissenschaftseigene, öffentlich geförderte Plattform EUDAT mit verschiedenen Datendiensten in Betrieb.¹⁹

Die digitale Archivierung ist mit Aufgabenstellungen großer Datenzentren in Vielem verwandt, zielt aber von vornherein auf andere Zeiträume sowie eine – idealerweise dauerhafte – Sicherung des digitalen Archivgutes, die eine Verfügbarkeit für teils auch noch unbekanntere Nutzungsformen erhält. Förderprojekte der letzten Jahre haben gezeigt, dass Langzeitarchivierung schwierige Fragestellungen beinhaltet – u. a. was die Nachverfolgung von Nutzungsergebnissen der Daten betrifft. Einige Grundprobleme der Langzeitarchivierung von digitalen Artefakten (Born Digitals oder Digitalisaten) sind derzeit technisch, logistisch und organisatorisch erst im Ansatz lösbar, da auf sich ändernder Hardware Forschungsdaten sowie Metadaten²⁰ gemeinsam mit den (rasch veraltenden) Programmen zu ihrer Nutzung archiviert werden müssen. Zudem sind organisatorische und rechtliche Standards digitaler Archivierung vielfach uneinheitlich bzw. unklar (etwa wo Behörden digitales Archivgut produzieren). Zur Illustration der Bandbreite der deutschen Institutionen, die sich mit digitaler Archivierung befassen, sind beispielsweise die Deutsche Nationalbibliothek, die Bayerische Staatsbibliothek, die Zentralen Fachbibliotheken, das Deutsche

Datenarchivierung,
Langzeitverfügbarkeit

¹⁸ Academia.edu – <https://www.academia.edu>; Mendeley Data – <https://data.mendeley.com>; ResearchGate – <https://www.researchgate.net> (zuletzt geprüft am 25.04.2016).

¹⁹ GBIF – <http://www.gbif.org>; NoMaD – <https://nomad-coe.eu>; ELIXIR – <https://www.elixir-europe.org>; Europeana – <http://www.europeana.eu>; EUDAT – <https://www.eudat.eu/what-eudat> (zuletzt geprüft am 25.04.2016).

²⁰ Vgl. Begriffsklärung im Anhang A.

Literaturarchiv Marbach, das Bundesarchiv sowie Landesarchive mit verschiedenen Kooperationsprojekten zu nennen.²¹ Wo Digitalisate gemeinsam mit Daten physisch herkömmlicher Art („originale“ Sammlungsgüter, Stoffproben, Grabungsfunde, Bohrkerne) archiviert werden, sind infrastrukturell ebenfalls besondere Herausforderungen gegeben.

Neben den institutionellen Akteuren wird das Bild durch eine Vielzahl von wissenschaftseigenen und kommerziellen, lokalen oder überregionalen bis hin zu internationalen *temporären Initiativen* geprägt, die teils *ad hoc*, teils mit mittelfristiger Perspektive, Forschungsdaten transferieren, aufbewahren, distribuieren und archivieren. Unter diesen Bedingungen ist das Schicksal von Forschungsdaten zurzeit teilweise schlichtweg dem Zufall überlassen. Die Projektförderung durch zeitlich befristete Drittmittel überwiegt gleichwohl nicht nur bei solchen wissenschaftsgetriebenen Aktivitäten, sondern auch dort, wo Forschungsdatenmanagement in institutionellem Rahmen vorangetrieben wird, also in Forschungsdaten- bzw. Informationszentren, Repositorien, Rechenzentren, Bibliotheken, Archiven oder Sammlungen.

Von den Aufgaben her gleichen sich die Arbeitsprofile von digitalen Informationsinfrastrukturservices unterschiedlicher Herkunft an. Hohe Interoperabilität muss hier das Ziel sein, ist gleichwohl aber nicht durchweg gegeben. In den genannten Feldern und Institutionen werden deshalb neue Formen der wissenschafts- und IT-nahen Datenkompetenz gebraucht. Fragen der Qualität, Produktivität und Integrität der Forschungsprozesse wie auch der Nachnutzbarkeit (oder ggf. Reproduktion) von Daten bedürfen arbeitsteilig-kooperativer Beurteilung. Für das Generieren von fachlichen und technischen Metadaten gilt Gleiches. Über Umfang und Art des Personalbedarfs für das umschriebene Feld – von der *digitalen* Bibliotheksexpertise über die *Kuratierung* von Forschungsdaten bis zu informatischen und fachmethodischen Schnittstellenkompetenzen – wird daher seit Jahren intensiv diskutiert.²²

2.3 TRÄGERSCHAFT – FINANZIERUNG – FÖRDERPROGRAMME

Trägerschaft und Finanzierung sind in der föderalen Wissenschaftslandschaft ähnlich vielfältig wie die Ausprägung der Dateninfrastrukturen selbst. Versucht man, einen Überblick zu gewinnen, muss zunächst gesagt werden, dass gesonderte Einrichtungen für das Forschungsdatenmanagement die Ausnahme sind und nicht die Regel. In verteilter Form existiert vielmehr ein breiter ‚Sockel‘, in dem Datenverantwortung schlicht bei den Forschungseinrichtungen oder bei

²¹ Zudem haben sich rund 20 Einrichtungen, die sich mit Archivierung befassen, im nestor Kompetenznetzwerk zur digitalen Langzeitarchivierung zusammengeschlossen: <http://www.langzeitarchivierung.de> (zuletzt geprüft am 25.04.2016).

²² Zum Verständnis der Begriffe Interoperabilität, Metadaten und Datenkuratierung siehe auch Begriffsklärungen im Anhang A.

den Forschenden selbst liegt. So werden Forschungsdaten – und nicht immer in kleineren Mengen – zunächst überall dort aufbewahrt, wo sie entstehen: in lokalen IT-Lösungen, an Standorten von Großgeräten, in Rechenzentren, in Bibliotheken und Bibliotheksverbänden, Archiven, aber auch bei Vertragspartnern öffentlicher Forschung (Verlage, IT-Unternehmen, Medienpartner, Stiftungen etc.) oder in der Cloud.

Die etablierten deutschen Forschungsdatenrepositorien sind oft disziplinär oder fachlich orientiert. Sie sind häufig im außeruniversitären Bereich angesiedelt und werden überwiegend kollaborativ und nicht selten mit internationalen Partnern betrieben. Finanzierungen erfolgen über die Trägereinrichtungen hinaus häufig über die Deutsche Forschungsgemeinschaft, das BMBF und die Forschungsrahmenprogramme der EU. Dies zeigt eine Auswertung der internationalen Datenbank re3data.org (siehe Anhang D.1).

Den ca. 220 bei re3data.org verzeichneten Diensten mit deutscher Beteiligung steht in einer die gesamte institutionelle Landschaft einbeziehenden Betrachtungsperspektive der eingangs angedeutete breite ‚Sockel‘ gegenüber. Für diesen stehen folgende Zahlen: 110 deutsche Universitäten sowie 233 Fachhochschulen mit jeweils teils relevant großen, teils kleineren datenhaltenden zentralen Einrichtungen, rund 280 Einrichtungen der außeruniversitären Forschung, mehrere Tausend wissenschaftliche bzw. wissenschaftsrelevante Bibliotheken, Archive und Museen – mit jeweils eigenen oder durch externe Partner betriebenen Repositorien. Hinzu kommen die Ressortforschungs- und Fachinformationseinrichtungen des Bundes sowie eine heterogene Gruppe von datenhaltenden Einrichtungen der Bundes- und Landesbehörden. Hier seien nur beispielhaft zu nennen: neun Landesrundfunkanstalten, allein knapp hundert Rechenzentren des Bundes und eine Vielzahl von Ämtern, deren millionenfache Datensätze bisher nur in geringem Umfang offen zugänglich sind.²³

Forschungsdateninfrastrukturen sind angesiedelt an Einrichtungen auf allen Ebenen des Wissenschaftssystems, werden also durch die Länder (Hochschulen, Museen), ansonsten überwiegend durch die von Bund und Ländern gemeinsam finanzierten Wissenschaftsorganisationen getragen. Hinzu kommen Bundes- und Länderverwaltungen. Auch die Kommunen und Unternehmen sind als Träger (Archive, Museen) involviert.

**Gesamtsystem:
viele Ebenen,
Aufgabenteilung
und Kollaboration**

²³ Die genannten Zahlen entstammen einer Erhebung der RfII-Geschäftsstelle vom Februar 2016. Sowohl eine Studie des BMI als auch die Expertenkommission Forschung und Innovation (EFI) kommen zu dem Ergebnis, dass die Potenziale von Open Government Data in Deutschland noch erheblich ausbaufähig sind. Vgl. EFI – Expertenkommission Forschung und Innovation (2016) – Jahresgutachten; Klessmann et al. (2012) – Open Government Data.

Die Aufgaben innerhalb des Datenlebenszyklus – Erhebung, Erschließung, Speicherung, Analyse, Archivierung, Zugang – werden nicht selten arbeitsteilig in Verbundstrukturen wahrgenommen. Der vorläufigen Analyse des RfII zufolge sind besonders die Hochschulen und die außeruniversitären Forschungseinrichtungen im Rahmen von Partnerschaften aktiv, im Bereich der Trägerschaft von Datenrepositorien bildet sich also das Stereotyp einer Trennung zwischen *Infrastruktureinrichtungen* als Anbietern und *Hochschulen* (oder *Forschungseinrichtungen*) als Nutzern nicht unbedingt ab (siehe Anhang D.1). Umgekehrt weisen z. B. die in der Leibniz-Gemeinschaft beheimateten Informationseinrichtungen oder auch datenhaltende Bundesbehörden in den letzten Jahren eine stärkere Forschungsorientierung auf.²⁴

Insbesondere für neue Informationsinfrastrukturen sind vor dem Hintergrund sich zunehmend verschränkender Zusammenarbeit zwischen Rechenzentren als Orten des ‚Rechnens‘ und Bibliotheken als Orten der Medienversorgung langjährige Aufbauphasen bis zur Reifung nicht ungewöhnlich. Anders als bei physischen Forschungsinfrastrukturen erfolgt die Implementierung digitaler Dienste nicht durch einmalige Maßnahmen wie *Gründung* oder *Bau*, sondern über Beta- und Erprobungsphasen, die – getragen von der Resonanz und dem Engagement einer mehr oder weniger schnell wachsenden Nutzercommunity – schrittweise Weiterentwicklungen nach sich ziehen. Viele Datenbanken, Softwaretools, Analyseplattformen und ähnliche Dienste sind aus Forschungsprojekten entstanden oder wurden als Projekte initiiert, z. B. durch Infrastruktur-Förderprogramme der DFG oder des BMBF. So können für verschiedene Entwicklungsstufen über zehn oder mehr Jahre Drittmittel eingeworben werden.

Die Start-up-Phase eines sich etablierenden Dienstes im Bereich des Forschungsdatenmanagements orientiert sich an den Belangen – und auch den Finanzierungsbedingungen – formal oder inhaltlich innovativer Datenproduktion. Prägend ist ein hoher Anteil an Forschungs- und Entwicklungsaktivitäten, die den Bedarfen eines Gesamtprojektes folgen. Typischerweise werden hier Forschungs- und Entwicklungsarbeiten u. a. von Doktoranden und jungen Post-docs durchgeführt. Ergebnisse sind publizierbar und tragen zur Profilbildung des wissenschaftlichen Nachwuchses bei.

Soweit sich die Projektinfrastrukturen später öffnen und ihre „Produkte“ einem breiteren Publikum anbieten, folgt eine Konsolidierungsphase, der allerdings keine Finanzierung mehr entspricht: Eine steigende Zahl von Forschenden außerhalb des Projektes nutzt, was da ist. Forschungsdateninfrastrukturen für solche Nutzungsformen zu öffnen, hat gewöhnlich einen positiven Effekt

²⁴ Es scheint nicht zuletzt die entschlossen Forschungsorientierung einfordernde Evaluationsstätigkeit des Wissenschaftsrates zu sein, die hier wirksam wird.

auf deren Qualität: Probleme und Fehler werden von Nutzern rasch erkannt, und im Idealfall erfolgt kurzfristig eine Rückmeldung an Administratoren und Entwickler, die zu einer Anpassung bzw. Verbesserung führt. Weiterhin fließen durch den erweiterten Nutzerkreis neue Ideen in die Entwicklung ein.

Gewinnen Dienste auf diese Weise Reputation in der Forschungslandschaft, wird von Seiten der Nutzer nicht nur eine hohe Verfügbarkeit²⁵ erwartet, sondern auch Beratung und Support bei Problemen. Der Betrieb spaltet sich in einen Forschungsbereich, der vor allem die Weiterentwicklung der Infrastruktur vorantreibt, und einen Servicebereich (Helpdesk und Feedback). Derlei Daueraufgaben im Bereich wissenschaftlicher Services sind ungeeignet für den wissenschaftlichen Nachwuchs, daher wächst dort der Bedarf für anderweitiges, allerdings hochqualifiziertes Personal. Spätestens ab diesem Zeitpunkt müsste die Verstetigung des Betriebes theoretisch durch institutionelle Mittel erfolgen. Jedoch zeigt die Erfahrung, dass Hochschulen oder Forschungseinrichtungen jenseits dessen, was das vorhandene Rechenzentrum mit geringem Mehraufwand übernehmen kann, nur selten darauf eingestellt sind, dies zu tun. Drittmittel für den Betrieb der Dienste lassen sich aufgrund der Förderstrukturen kaum einwerben. Drittmittel für die Weiterentwicklung digitaler Infrastrukturen können allerdings beantragt werden. Dies birgt die Gefahr, dass bestehende Dienste zu Forschungsprojekten umetikettiert werden, oder aber das Rad mehrfach neu erfunden wird. Eine Fehlallokation von Forschungsmitteln kaschiert dann ein Infrastrukturproblem.

Bei aus Projekten entstandenen Infrastrukturen, denen besondere gesamtstaatliche Bedeutung zugemessen wird, ist in Einzelfällen nach Evaluation eine Aufnahme in die gemeinsame Forschungsförderung von Bund und Ländern erfolgt. Beispiele hierfür sind die nationalen Längsschnittstudien NEPS und SOEP.²⁶ Große Dateninfrastrukturen der außeruniversitären Forschungseinrichtungen werden von Bund und/oder Ländern teils direkt etabliert – ein Beispiel ist das Deutsche Klimarechenzentrum (DKRZ) –, teils wird ihre Etablierung indirekt über die Haushalte der Trägereinrichtungen finanziert, wie z. B. die Datenzentren der Max-Planck-Gesellschaft.

Schwierige Finanzierung von Services

²⁵ Zum Verständnis des Begriffs Verfügbarkeit siehe auch ausführliche Begriffsklärung im Anhang A.

²⁶ Die Längsschnittstudie NEPS – National Education Panel Study – wurde von 2009 bis 2013 durch das BMBF gefördert. Die Verstetigung erfolgte nach Evaluation durch den Wissenschaftsrat mit der Gründung des heutigen Leibniz-Instituts für Bildungsverläufe (IfBi). Vgl. GWK (2013) – Ergebnisse der Sitzung (Pressemitteilung), S. 2; Evaluation Leibniz-Institut für Bildungsverläufe: <http://www.wissenschaftsrat.de/nc/arbeitsbereiche-arbeitsprogramm/evaluation.html#c20161> (zuletzt geprüft am 25.04.2016).

Das „Sozio-oekonomische Panel“ (SOEP) erhielt von 1983 bis 2002 eine Projektförderung durch die DFG. Seit 2003 wird es als „Serviceeinrichtung“ zu zwei Dritteln vom Bund und zu einem Drittel vom Land Berlin gefördert und ist am Deutschen Institut für Wirtschaftsforschung angesiedelt, einem Institut der Leibniz-Gemeinschaft. Vgl. http://www.diw.de/de/diw_02.c.221178.de/ueber_uns.html (zuletzt geprüft am 25.04.2015).

Kaum Verstetigungs- optionen

Der Ausbau von Forschungsdatenservices oder die Verstetigung von Projekten sind den wissenschaftlichen Partnern und ihren Trägereinrichtungen weitgehend selbst überlassen. Was hier gelingt, ist in hohem Maße von der Bereitschaft und den Möglichkeiten der Trägereinrichtung abhängig. Dass wissenschaftliche Dienstleistungen unter diesen Bedingungen nachhaltig übergreifend angeboten werden, setzt entweder die Angliederung an einen traditionellen, bundesweit tätigen Akteur voraus, z. B. aus dem Bibliotheks- und Archivsystem, aus der Leibniz- oder Helmholtz-Gemeinschaft oder anderen außeruniversitären Wissenschaftseinrichtungen. Oder es werden neue Formen der Kooperation geschaffen, z. B. zwischen Hochschulen auf Länderebene oder zwischen Hochschulen und außeruniversitären Forschungseinrichtungen; weitere Beispiele für eine stabile Selbstorganisation sind der DFN-Verein oder (wobei der Wissenschaftsrat hier Potenzial für eine strukturelle Fortentwicklung sieht)²⁷ die Gauß-Allianz.

Im föderalen System der Bundesrepublik ist der Manövrierspielraum für den Aufbau von länder- und sektorenübergreifend gemeinsam getragenen und genutzten Forschungsdateninfrastrukturen äußerst eng. Bottom-up-Prozesse – für die es an Anreizen nicht mangelt – führen daher aktuell vielfach trotz großen Engagements zu kurzlebigen Lösungen und manchmal in Sackgassen. Das Bild wirkt diffus und nicht nachhaltig. Knappe Ressourcen kommen ineffizient zum Einsatz. Auch sind im Falle einer institutionellen Verstetigung die Formen, Akteure und Kriterien von Evaluationen uneinheitlich.

Heterogene Förderlandkarte für Informations- infrastrukturen

Viele Forschungsförderer haben den Bedarf an Forschung und Entwicklung im Bereich digitaler, datenintensiver Methoden und fachwissenschaftlich maßgeschneiderter Werkzeuge erkannt. Seit Jahren werden im Rahmen von Projektförderung innovative Ansätze nicht nur in der forschenden Informatik, sondern auch – im Sinne von hilfswissenschaftlichen, infrastrukturnahen Lösungen – an interdisziplinären Schnittstellen von disziplinärer und IT-bezogener Forschung unterstützt. Dabei werden zunehmend Aussagen über *Nachhaltigkeit* oder Zusagen zum Beispiel im Sinne einer Verstetigung gefordert.

Dies ist namentlich bei der Beantragung von dezidierten Infrastrukturmitteln mittlerweile der Fall. Wenn allerdings beispielsweise Mittel des Sitzlandes einer Universität oder einer Forschungseinrichtung eingesetzt werden sollen, kann sich in der Praxis die Verstetigung von gemeinsam getragenen und genutzten Angeboten kompliziert gestalten. Ebenso sind befristete Projekte mehrerer – aber nicht aller Bundesländer – nur schwer in einer späteren Phase im Sinne einer bundesweiten Lösung ‚auszurollen‘. Anreize der länderübergreifenden Projektförderung (etwa durch DFG oder BMBF) führen daher trotz engagierter Akteure selten zu langlebigen Lösungen.

Die so entstandene Förderlandkarte bleibt instabil und heterogen. Zur Bewahrung der in Projekten aufgebauten wissenschaftlichen Ressourcen und zum

²⁷ WR (2015) – Empfehlungen zur Finanzierung des NHR.

Aufbau bundesweiter wissenschaftlicher Informationsinfrastrukturen sind daher in den vergangenen Jahren erste koordinierende Aktivitäten entstanden.²⁸ Ein neues Steuerungsinstrument zur Entwicklung groß angelegter wissenschaftlicher Informationsinfrastrukturen ist die 2013 erstmals veröffentlichte nationale Roadmap für Forschungsinfrastrukturen des BMBF (siehe auch 2.4).

Was bleibt, ist eine große Zuständigkeitslücke für die zielgerichtete Gestaltung und den Ausbau des Gesamtsystems, die gerade den *mittleren* Bereich von fachnahen, thematischen oder auch netzwerkartigen Infrastrukturen betrifft. Förderprogramme, die in erster Linie inhaltliche Projekte fördern wollen und sollen, sind hierauf nicht nur in finanzieller Hinsicht nicht eingestellt. Auch Begutachtungen und Weiterentwicklungen, die während einer Aufbauphase noch gewährleistet sind, lassen sich in der Breite des Wissenschaftssystems nicht über die üblichen Förderformate hinaus implementieren.

Fehlende Zuständigkeit für den Ausbau des Gesamtsystems

Die Ermittlung eines Gesamtbedarfs an Finanzierung und Investition für wissenschaftliche Informationsinfrastrukturen bezeichnen die wissenschaftlichen und die wissenschaftspolitischen Akteure in Deutschland und anderen Ländern aufgrund der Vielfalt der in Betracht zu ziehenden Aspekte übereinstimmend als schwierig. Jedenfalls besteht, das ist Konsens, erheblicher Finanzierungs- und Investitionsbedarf auf verschiedenen Ebenen des Gesamtsystems.

Erheblicher Finanzierungs- und Investitionsbedarf

Die reinen Unterhaltskosten für den Betrieb einer Datenbank sind aus Erfahrungswerten gut abschätzbar. Klar ist, dass die vorhandene Vielfalt Ressourcen bindet, und dass bei entsprechenden Anfangsinvestitionen Synergiepotenziale bestehen. Im Bereich der Datenhaltung sind große Infrastrukturen durch Skaleneffekte effizienter zu betreiben als kleine. Freilich ist mit einer stark steigenden Zunahme von Daten und Datennutzung zu rechnen, so dass eine gut integrierte Forschungsdatenlandschaft in Deutschland in der Summe nicht Kosten sparen, sondern die Leistungsfähigkeit des Gesamtsystems Wissenschaft erhöhen wird. Dies ist für die internationale Wettbewerbsfähigkeit in dringlicher Weise nötig.

Ein vielfach unterschätzter und gleichwohl hoher Bedarf²⁹ besteht außerdem an Ressourcen für qualifiziertes Personal, das es heute – auch was Ausbildungswege angeht – noch gar nicht gibt. Es sind aber die ‚Köpfe‘, die durch Integration von Daten, Informationen und Wissen aus heterogenen Quellen und über Domänengrenzen hinweg für die Entstehung neuen wissenschaftli-

Erheblicher Personalbedarf

²⁸ Zu nennen sind hier das von der DFG koordinierte System von Fachinformationsdiensten und das Förderkonzept Medizininformatik des BMBF (vgl. Anhang D.4), ebenso Gremien wie der RatSWD (www.ratswd.de) oder Projekte wie GFBio (www.gfbio.org).

²⁹ Schätzungen gehen für die EU von mehreren Hunderttausend Stellen aus. Vgl. Interview vom 09.03.2016 mit Barend Mons, Vorsitzender der High Level Expert Group for the European Open Science Cloud, <http://primeurmagazine.com/weekly/AE-PR-05-16-58.html> (zuletzt geprüft am 25.04.2016).

chen Wissens und damit für die Wertschöpfung aus den Daten sorgen.³⁰ Erste Förderlinien setzen bereits die Nutzung digitaler Informationsinfrastrukturen als Förderziel.³¹ Dies ist im Sinne der nachhaltigen Nutzung von digitalen Ressourcen und des dringend erforderlichen Kompetenzaufbaus sehr zu begrüßen.

Schlecht geklärt und technisch wie organisatorisch aufwändig ist überdies die Bewältigung der Archivierung und Langzeitverfügbarkeit von Daten und Datendiensten. Eine Aufbewahrung von digitalen wissenschaftlichen Sammlungen zeitstabil und potenziell auf Dauer zu gewährleisten, stellt für die Wissenschaftsfinanzierung eine große Herausforderung dar. Unter den Bedingungen einer vorwiegend projektförmigen Finanzierung³² wird sie nicht zu bewältigen sein.

2.4 DEUTSCHLAND IM INTERNATIONALEN VERGLEICH

Der weltweite Zugang zu wissenschaftlicher Information hat sich durch das Internet in den letzten 15 Jahren grundlegend verändert. Hierbei formen Bedarfe der Information und Bildung, nicht zuletzt aber auch wirtschafts- und innovationspolitische Trends den Hintergrund für die wissenschaftspolitische Diskussion. Offenen Zugang zu Information fordern Bürgerbewegungen, den offenen Zugang speziell zu Daten postulieren Studien und Empfehlungen der OECD und der G7/G8-Staaten.³³ Parallel dazu warnt ein Bericht der UNO vor der international wachsenden Ungleichheit zwischen „data-rich“ und „data-poor countries“ sowie vor einer „Unsichtbarkeit“ bestimmter Gruppen von Menschen im Hinblick auf ihre Daten.³⁴

Die internationale Landschaft ist allerdings nicht nur im Hinblick auf den Umgang von Wissenschaft und Infrastrukturanbietern mit Forschungsdaten von Ungleichheiten und Ungleichzeitigkeiten geprägt. Sie ist auch in puncto Strategien und Entwicklungspfade bis hin zu *Kulturfragen* heterogen.³⁵ Dies prägt die Gestaltung wissenschaftlicher Kooperationen in erheblichem Maße und bringt auch aus deutscher Perspektive die Notwendigkeit des Abgleichs und der Positionierung mit sich.

Ungleichheit und
Ungleichzeitigkeit
der internationalen
Entwicklung

³⁰ Vgl. hierzu auch die Begriffsklärung zur dynamischen Wissensintegration im Anhang A.

³¹ Für Beispiele siehe Anhang D.4, Übersicht Förderprogramme.

³² Das Kompetenznetzwerk nestor hat in seinem Positionspapier 2015 darauf hingewiesen, dass auch die Langzeitarchivierung derzeit überwiegend projektförmig finanziert wird. Vgl. Nestor (2015) – Positionspapier Digitale Langzeitarchivierung.

³³ Allianz der Wissenschaftsorganisationen (2003) – Berliner Erklärung; Chan et al. (2002) – Budapest Open Access Initiative (Webseite); G8 (2013) – Open Data Charter; OECD (2008) – Recommendation Public Sector Information.

³⁴ IEAG (2014) – A world that counts, S. 3 u. 7.

³⁵ Diese Sicht repräsentiert die Studie der OECD (2015) – Making Open Science a Reality.

Der auf der operativen Ebene notwendige Aufbau verteilter internationaler Informationsinfrastrukturen wird auf der wissenschaftspolitischen Ebene mittels Roadmaps vorangebracht bzw. flankiert, an denen sich Deutschland beteiligt. Auf der Ebene der wissenschaftlichen Prozesse ist darüber hinaus die Entwicklung von Standards und Best Practices in internationalen Netzwerken relevant.

Zur Entwicklung nationaler Informationsinfrastrukturen sind unter anderem Policy-Initiativen von Interesse. International fällt ein deutlich heterogenes Bild zum Forschungsdatenmanagement (FDM) auf, das sich zudem stark im Wandel befindet.³⁶ Es spiegeln sich darin die unterschiedlichen Ausgangslagen der Länder wider, u. a. Größe, Wirtschaftskraft, föderale oder zentrale Strukturen der staatlichen Organisation, kulturelle Identität. Entsprechend ist das Wechselspiel von Bottom-up- und Top-down-Aktivitäten in den jeweiligen Ländern unterschiedlich ausgeprägt.

Beispielsweise wird im föderalen, großflächigen Australien „Egalität“ gelebt, Bottom-up-Aktivitäten zur Harmonisierung einzelner FDM-Themen stehen nationalen Top-down-Policy-Initiativen gegenüber, die eine zentrale Komponente des wirtschaftlichen Strukturwandels darstellen. Über die National Collaborative Research Infrastructure Strategy (NCRIS) wurde 2009 mit dem Australian National Data Service (ANDS) ein Projekt initiiert, das brückenbildende Aufgaben in der Zusammenführung von Bottom-up- und Top-down-Aktivitäten übernimmt. Ein wesentliches Element der Wirksamkeit sind die längerfristig angelegten Projektlaufzeiten (z. B. 2 x 10 Jahre im Rahmen der NCRIS-Förderung). Erfolgreiche Initiativen werden von ANDS besonders unterstützt, um eine Sogwirkung zu erzielen.

Auch Kanada weist eine stark dezentrale staatliche Organisationsstruktur auf. Die Entwicklung des FDM basierte bis dato hauptsächlich auf Graswurzelinitiativen, die durch ihre sehr hohe Nachfrage punktuell eine Top-down-Koordination auslösten. Beispielsweise wurde von der Canadian Association of Research Libraries die Portage Initiative gegründet, die Unterstützung in FDM bereitstellt. Diskussionsforen wie Research Data Canada und der Leadership Council for Digital Infrastructure versuchen ohne finanzielle Unterstützung auf stärkere Top-down-Strukturen im FDM hinzuwirken.

Ein völlig anderes Beispiel stellen die Niederlande dar. Top-down-Strukturen in der FDM- und Open-Access-Strategie sind etabliert. Ziele sind zum einen der Nachweis von Daten, die Förderung von Transparenz sowie die Qualitätssicherung im Sinne guter wissenschaftlicher Praxis, zum anderen sollen Datensammlungen aufgebaut werden. In den Niederlanden fallen besonders die auf Kooperation mehrerer Akteure angelegten nationalen Lösungen auf, z. B. die

Internationale Policy-Initiativen für das Forschungsdatenmanagement

Australien und Kanada: Bottom-up getriebene Strategien

Niederlande: Kollaborative Strukturen

³⁶ Unmittelbare Eindrücke wurden in einer Experten-Konsultation der AG Internationale Orientierung vom April 2016 gewonnen, die genannten Länderbeispiele erheben keinen Anspruch auf Vollständigkeit.

Research Data Netherlands im Bereich der Wissenschaft und die sektorübergreifende Netherlands Coalition for Digital Preservation (NCDD) im Bereich der Archivierung. Bereits 2005 wurde DANS als gemeinsame Serviceeinrichtung der Königlich-Akademie für Kunst und Wissenschaft (KNAW) und der Netherlands Organisation for Scientific Research (NWO) gegründet, um den Zugang zu digitalen Forschungsdaten dauerhaft zu ermöglichen. Bei der Umsetzung der nationalen FDM-Strategie werden die Forschenden direkt eingebunden. Beispielsweise wird in einem Pilotprojekt der NWO seit Januar 2015 die Bewilligung zur Forschungsförderung von einem Datenmanagementplan nach erfolgreicher wissenschaftlicher Projektbegutachtung abhängig gemacht. Die Erkenntnisse aus diesem Pilotprojekt sollen in einer umfassenden Policy zu Open Access zu Daten Eingang finden. Die Regierung verfolgt außerdem die Vision eines „Open Government“, die Open Data, E-Government und Bürgerbeteiligung einschließt.³⁷

USA und EU:
Datenmanagement
zunehmend
verpflichtend

Die US-amerikanische National Science Foundation (NSF) verlangt, anders als in den Niederlanden, die Vorlage von Datenmanagementplänen bereits bei der Projektbeantragung. Die US-amerikanischen National Institutes of Health (NIH) verpflichten ihre Drittmittelnehmer schon seit 2007 auf die Veröffentlichung der Ergebnisse sowohl in Form von Publikationen als auch in Gestalt des Zugänglichmachens von Daten. Im aktuellen Forschungsrahmenprogramm der EU wird die verpflichtende Veröffentlichung von Forschungsdaten im Rahmen eines „Open Data Pilot“ erprobt. Die Zwischenergebnisse lassen eine Ausweitung auf weitere Teilnehmer am Forschungsrahmenprogramm erwarten. Für die Projekte, die hier teilnehmen, sind Datenmanagementpläne Pflicht.³⁸

Konvergenz, zentrale
Punkte

Insgesamt lässt sich international eine bemerkenswerte Konvergenz in Richtung auf die folgenden zentralen Punkte beobachten:

- Es gibt vielfältige Aktivitäten, die auf die Entwicklung von FDM-Strategien bzw. -plänen (Roadmaps) abzielen. Als wesentliches Element erscheint eine Ergänzung der Bottom-up-Aktivitäten durch nationale Strategien und weitere Instrumente. Diese sind in einzelnen Ländern noch sehr unterschiedlich ausgeprägt.
- Datenmanagementpläne werden als Instrument einer Anreizsetzung zur Motivation der Forschenden als sinnvoll erachtet.
- Nachhaltigkeit ist als das zentrale Problem allerorts erkannt und wird teils durch langfristige Förderinstrumente angegangen.

³⁷ Research Data Netherlands – <http://www.researchdata.nl>; NCDD – <http://www.ncdd.nl/en/about-the-ncdd>; DANS – <http://www.dans.knaw.nl>; Open Government Partnership Netherlands – <http://www.opengovpartnership.org/country/netherlands> (zuletzt geprüft am 25.04.2016).

³⁸ European Commission- DG for Research and Innovation (2016) – Guidelines on Data Management.

- Die Notwendigkeit eines Datenqualitätsmanagements (Data Quality Controlling) wird gesehen, messbare Indikatoren fehlen bislang, Evaluierungen sind ansatzweise vorhanden.
- Die Notwendigkeit der Entwicklung von Ausbildungs- und Trainingsprogrammen (Studiengängen) ist erkannt.

In Deutschland sind ähnliche Entwicklungen zu beobachten. Das Forschungsdatenmanagement hat zunächst über die Aktivitäten der Allianz-Initiative (ab 2010), dann über die Empfehlungen der HRK und des Wissenschaftsrats (2012) sowie die Aufnahme in die Digitale Agenda der Bundesregierung Fahrt aufgenommen.³⁹ Mit verschiedenen Landesinitiativen, der Helmholtz Data Federation, den Leitlinien der DFG und einigen Förderrichtlinien sind in der Folge konkrete Entwicklungen angestoßen worden, um das Forschungsdatenmanagement institutionell stärker zu verankern. Ein flächendeckendes Angebot an Informationsinfrastrukturen für das Forschungsdatenmanagement existiert jedoch ebenso wenig wie die Formulierung einer nationalen Strategie.

Deutschland:
Zahlreiche
Einzelinitiativen

In der EU findet seit 2002 eine strategische Entwicklung von transeuropäischen Forschungsinfrastrukturen über die Roadmap des European Strategy Forum on Research Infrastructures (ESFRI) statt.⁴⁰ Die 2006 erstmals veröffentlichte ESFRI-Roadmap ist zugleich ein wichtiges Instrument für die Konsolidierung der Infrastrukturen und der mit ihnen verbundenen Prozesse. Sie wird über Maßnahmen in den EU-Forschungsrahmenprogrammen finanziell unterstützt, die Infrastrukturen auf der Roadmap werden aber im Wesentlichen durch die Mitgliedstaaten und assoziierten Länder über das Prinzip der variablen Geometrie finanziert. Die Geschichte der ESFRI-Projekte zeigt gut auf, welch langer Atem im Bereich der Infrastrukturentwicklung notwendig ist. Von den 48 Projekten der Roadmaps 2001-2010 galten 2013 zehn als „Erfolgsgeschichten“, 17 weitere haben seitdem die Implementierungsphase erreicht und werden in der Roadmap 2016 als „Landmarks“ bezeichnet. Die ESFRI-Evaluation von 2013 hat gezeigt, dass die meisten Forschungsinfrastrukturen auch nach mehreren Jahren Entwicklungszeit weiterhin „substanzielle Unterstützung und Führung“ benötigten. Dies betrifft sowohl das Management der Infrastruktur, die Stabilität ihrer Finanzierung als auch das Engagement ihrer Stakeholder. Als ebenso ausbaufähig bewertet wurden darüber hinaus das Projektmanagement, die Nutzerstrategie und die Risikoevaluation. Auch ethische Aspekte seien nicht ausreichend behandelt worden.⁴¹ Die typischen Nachhaltigkeitsprobleme projektförmig geförderter Infrastrukturen sollen über eine Reihe von Anpassungen gelöst werden. So werden für die Aufnahme in die neue ESFRI-Roadmap ab 2016 nachhaltige Förderzusagen der beteiligten Staaten und ein höherer Reifegrad der Infrastruktur vorausgesetzt.

ESFRI:
International verteilte
Informationsinfra-
strukturen für die EU

³⁹ Vgl. Übersicht zur Historie im Anhang B.

⁴⁰ Vgl. <http://www.esfri.eu/about-esfri> (zuletzt geprüft am 25.04.2016).

⁴¹ European Commission (2013) – Assessment ESFRI projects.

Der ESFRI-Roadmap-Prozess wurde nach und nach mit Roadmap-Prozessen auf nationalstaatlicher Ebene verschränkt.⁴² Deutschland hat als Ergebnis eines Pilotprozesses erstmals 2013 eine „Roadmap für Forschungsinfrastrukturen“ veröffentlicht. Sie enthält 27 Projekte, von denen 16 zugleich auf der ESFRI-Roadmap zu finden sind.⁴³

Im August 2015 wurde in Deutschland basierend auf den Erfahrungen der Pilotphase der Nationale Roadmap-Prozess für Forschungsinfrastrukturen gestartet.⁴⁴ Er besteht aus zwei einander ergänzenden Bewertungsverfahren: Einer wissenschaftsgeleiteten Bewertung, die durch einen mandatierten Ausschuss des Wissenschaftsrates durchgeführt wird, und einer wirtschaftlichen Bewertung der zu erwartenden Kosten, der Projektreife und der Umsetzungschancen, die sich eng an die Controlling-Kriterien des BMBF für die Durchführung von Großprojekten anlehnt und durch den beauftragten Projektträger organisiert wird. Beide Bewertungen werden unter Einbeziehung externer internationaler Gutachter durchgeführt und ergeben voraussichtlich im Sommer 2017 ein Votum zur forschungspolitischen Priorisierung langfristig orientierter Forschungsinfrastruktur-Vorhaben. Der nationale Roadmap-Prozess für Forschungsinfrastrukturen ist kein Förderprogramm, sondern ein Verfahren zur Unterstützung und Fundierung strategischer forschungspolitischer Entscheidungen. Er ist mit einer grundsätzlichen Förderabsicht für die priorisierten Projekte hinterlegt, wobei eine Antragstellung auf Förderung erst nach erfolgreichem Durchlaufen des Roadmap-Verfahrens möglich ist (siehe Anhang D.2 für einen vergleichenden Überblick über Bewertungsdimensionen und Kriterien von ESFRI und BMBF). Die Entwicklung der Evaluationen durch die DFG (Kommission für IT-Infrastruktur und LIS – Wissenschaftliche Literaturversorgungs- und Informationssysteme) hat in den letzten 15 Jahren gezeigt, wie langsam die Evolution von Begutachtungsprozessen fortschreitet, wenn keine nationalen Roadmaps dies befördern.

Im Zuge der europäischen Roadmap-Prozesse und der internationalen Zusammenarbeit werden vernetzte Informationsinfrastrukturen geformt, die sich aus mehreren institutionellen oder nationalen „Knoten“ speisen. Verteilte Standorte vernetzen sich also nicht nur, sondern bringen sich in eine übergreifende internationale Infrastruktur ein. Die Europäische Kommission plant hier für die kommenden Jahre die Bildung einer „European Open Science Cloud“ sowie einer europäischen Dateninfrastruktur für den Bereich des Hochleistungsrech-

⁴² 22 Mitgliedstaaten haben in der Folge nationale Roadmaps implementiert. Vgl. European Commission (2015) – European Research Area. Facts and Figures 2014, S. 21. Für eine Übersicht des Standes in den Mitgliedsstaaten siehe <http://www.esfri.eu/national-roadmaps> (zuletzt geprüft am 25.04.2016).

⁴³ BMBF (2013) – Pilotprojekt nationale Roadmap; WR (2013) – Bewertung Forschungsinfrastrukturvorhaben.

⁴⁴ BMBF (2016) – Nationaler Roadmap-Prozess.

nens. Die Initiative soll die Entwicklung von Open Science und die Schaffung eines europäischen Digitalen Binnenmarktes befördern.⁴⁵

Verteilte Informationsinfrastrukturen werden mittlerweile als eigene Organisationsform angesehen, die einer besonderen Governance und auch Rechtsform bedürfen.⁴⁶ Auf europäischer Ebene wurde 2012 eine eigene Rechtsform geschaffen, das *European Research Infrastructure Consortium* (ERIC). ERICs sind rechtlich selbstständige Konstrukte mit definierten Governance-Strukturen, die Personal einstellen, Drittmittelnehmer sein können und von der Mehrwertsteuer befreit sind. Ihre Einrichtung muss durch die Europäische Kommission genehmigt werden.⁴⁷ In Deutschland können seit 2013 ERICs gegründet werden, wobei das ERIC-Framework primär das Zusammenwirken internationaler Partner regelt; ein ERIC kann folglich nicht als rein deutsche Unternehmung gegründet werden. Ein nationales Äquivalent gibt es nicht. So weisen beispielsweise Vertreter der Geistes- und Kulturwissenschaften darauf hin, dass es für die digitalen Forschungsinfrastrukturen spezifischer Organisationsformen bedarf, die noch zu entwickeln sind.⁴⁸

Besondere
Organisationsformen

Deutsche Beteiligungen an international verteilten Informationsinfrastrukturen bedürfen nicht selten einer Konsolidierung. Ein Beispiel hierfür ist das lebenswissenschaftliche Repositorium ELIXIR, dessen langwierige nationale Knotenbildung eine Beteiligung Deutschlands lange hinauszögerte.⁴⁹ Ein erfolgreiches Beispiel stellt hingegen das von 18 europäischen Staaten getragene Projekt DARIAH mit seinem nationalen Knoten DARIAH-DE dar.⁵⁰ Dort engagieren sich bereits 21 Projektpartner aus Deutschland; DARIAH hat den Rechtsstatus eines ERIC erreicht.

Insgesamt ist Deutschland im EU-Vergleich einer der sehr aktiven Akteure, sowohl beim Aufbau international vernetzter Forschungsinfrastrukturen als auch bei der Beteiligung an ihnen. So engagiert sich Deutschland finanziell an 18 der 48 ESFRI-Projekte und ist Sitzland mehrerer Forschungsinfrastrukturen.⁵¹ Allerdings sind die langen Entwicklungszeiten, die ungelösten Langzeitfinanzierungen und Akzeptanzprobleme in der Forschung ein nach wie vor bestehendes Problem der Förderung. Vergleichbare Initiativen gibt es in der globalen

⁴⁵ European Commission (2016) – Data and knowledge economy.

⁴⁶ GSF- Global Science Forum (2014) – IDRIS. Issues and Options.

⁴⁷ European Commission (2010) – ERIC Legal Framework.

⁴⁸ DARIAH-DE (2016) – Memorandum digitale Forschungsinfrastrukturen (Webseite).

⁴⁹ Eine Konsolidierung scheint jetzt – nach einer Vorlaufzeit von etwa zehn Jahren – mit dem de.NBI Netzwerk (Deutsches Netzwerk für Bioinformatik-Infrastruktur) gelungen, vgl. <https://www.elixir-europe.org/news/elixir-and-denbi-agree-collaboration-strategy> (zuletzt geprüft am 25.04.2016).

⁵⁰ DARIAH-DE – <https://de.dariah.eu>; DARIAH-EU – <https://dariah.eu> (zuletzt geprüft am 25.04.2016).

⁵¹ BMBF (2014) – Internationale Kooperation.

wissenschaftlichen Zusammenarbeit nur wenige.⁵² Als weiteres beispielhaftes Projekt wäre aus dem Bereich der Geo- und Erdsystemforschung das Global Earth Observation System of Systems (GEOSS)⁵³ zu nennen, in dem Deutschland Mitglied ist.

Die Tendenz zur grenzübergreifenden Vernetzung von Informationsinfrastrukturen erzeugt Bedarf nach Standardisierung und Harmonisierung. Unabhängig von den *top down* organisierten Roadmap-Aktivitäten haben sich international eine Vielzahl *bottom up* getriebener Best-Practices- und Standardisierungsinitiativen entwickelt. Diese haben Empfehlungen und technische Standards hervorgebracht, deren hohe Anzahl allerdings teilweise die Zersplitterung und Silo-Bildung innerhalb der internationalen Landschaft eher begünstigt als reduziert.⁵⁴ In dieser Lage bemüht sich die 2013 gegründete globale Research Data Alliance (RDA) – ebenfalls mit deutscher Beteiligung – um eine Konsolidierung.

Initiativen aus einzelnen EU-Ländern, wie das in den Niederlanden entwickelte *Data Seal of Approval* oder das deutsche *nestor*-Siegel für vertrauenswürdige Langzeitarchive, treffen auch international auf Interesse und sind entsprechend vernetzt (vgl. 2.5). Auch der RatSWD, der national sehr aktiv zur Entwicklung „guter Praxis“ bei den Forschungsdatenzentren beigetragen hat, verabschiedete 2014 eine Internationalisierungsstrategie. In der Breite sind aber nach bisherigem Eindruck die Angebote zum Erfahrungsaustausch und zur Mitgestaltung bei RDA und anderen Akteuren in Deutschland noch zu wenig bekannt. Mit RDA-DE⁵⁵ etabliert sich allerdings gerade ein Multiplikator auf nationaler Ebene.

Eine förderliche Wirkung dieser Initiativen auf den internationalen *Forschungsdatenmarkt* und seine Infrastrukturen wird aller Voraussicht nach von einer zuverlässigen Bereitstellung ausreichender Ressourcen abhängen. Derzeit werden sie durch ehrenamtliches Engagement von Expertinnen und Experten getragen, das auch zulasten von Forschung oder Pflichten in der jeweiligen Heimatorganisation erfolgt und wenig Anreiz bietet. Unregelmäßige Beteiligung und fehlende strategische Rückbindung in die Heimateinrichtungen stellen Risiken dar.

⁵² Allerdings diskutieren sowohl das OECD Global Science Forum als auch die Weltverbände der Forschungsförderer gemeinsame Initiativen zum Aufbau global vernetzter Forschungsinfrastrukturen.

⁵³ GEOSS – <http://www.earthobservations.org/geoss.php> (zuletzt geprüft am 25.04.2016).

⁵⁴ Siehe dazu auch Dally/Fless/Förtsch (2012) – *Altertumswissenschaften*.

⁵⁵ In RDA aktive Experten aus Deutschland organisieren Erfahrungsaustausch über Workshops – vgl. <http://www.forschungsdaten.org/index.php/RDA-DE> (zuletzt geprüft am 25.04.2016).

2.5 ERMÖGLICHUNGSSTRUKTUREN UND BESONDERE ENTWICKLUNGSAUFGABEN

Ein Aspekt des Konsolidierens von Datenmanagement-Initiativen ist die Schaffung von Orientierung. Entscheider im Wissenschaftssystem und Forschende sind darauf ebenso angewiesen wie die breitere Öffentlichkeit, Wirtschaft und Politik. Den Bedarf an besserer Auffindbarkeit und Identifizierbarkeit von Datenbeständen bedienen erste Register und Verzeichnisse, die auf die Belange der in Deutschland basierten Forschung zugeschnitten sind: Die Datenbank RIsources der DFG hält beispielsweise Informationen über rund 300 Forschungsinfrastrukturen vor (davon etwa zwei Drittel Informationsinfrastrukturen und Datenrepositorien). re3data.org, ein auf Datenrepositorien spezialisiertes Verzeichnis, listet knapp 250 international ausgerichtete Forschungsdatenrepositorien mit deutscher Beteiligung, es handelt sich um überwiegend kollaborative Strukturen (siehe auch Analyse in Anhang D.1). Das Informationszentrum GESIS hat mit DARA ein Register speziell für sozialwissenschaftliche Forschungsdaten etabliert. Verzeichnisse dieses Typs gehören derzeit zu den wenigen Quellen, denen quantitative Daten zur Entwicklung der kommenden deutschen Infrastrukturlandschaft zu entnehmen sind, sie sind also nicht zuletzt für Nutzer aus der Infrastrukturpolitik interessant. Das Potenzial dieser Registraturen wird allerdings bisher nur in einem kleinen Kreis wahrgenommen.⁵⁶

Verlässliche Identifikatoren (*persistent identifier*) sind die Voraussetzung für die Funktionalität solcher Register, die als *Ermöglichungsstrukturen* bezeichnet werden können. Identifikatoren gewährleisten Auffindbarkeit und Vernetzung einzelner veröffentlichter Datensätze, Personen und Organisationen. Dazu haben sich in den letzten Jahren weltweit operierende Dienste entwickelt.⁵⁷ Mit DataCite⁵⁸ existiert hier eine erfolgreiche Bottom-up-Initiative mit namhafter deutscher Beteiligung.

Auffindbarkeit durch Identifikatoren

Die derzeit herrschende *Gründerzeit* fördert das Entstehen zahlreicher, unterschiedlich ausgeprägter Dienste und Services für unterschiedliche Nutzergruppen. Erste wissenschaftseigene Standardisierungs- und Qualitätsmanagementinitiativen für Informationsinfrastrukturen bieten hier Potenzial für Orientierung und Professionalisierung. Der gewählte Weg ist derjenige des Angebotes einer freiwilligen, wissenschaftsintern organisierten Zertifizierung. Für Deutschland sind die Akkreditierung durch den RatSWD im Bereich Sozial- und Wirtschafts-

⁵⁶ DFG RIsources – <http://risources.dfg.de>; DARA – <http://www.da-ra.de>; re3data- <http://www.re3data.org> (zuletzt geprüft am 25.04.2016). Die Angaben zur Anzahl der verzeichneten Dienste stammen vom April 2016.

⁵⁷ Z.B. das Digital Object Identifier System (DOI) – <https://www.doi.org> und die ORCID-Initiative (Open Researcher and Contributor ID) – <http://orcid.org> (zuletzt geprüft am 25.04.2016).

⁵⁸ DataCite – <https://www.datacite.org> (zuletzt geprüft am 25.04.2016), die Geschäftsführung ist bei der TIB Hannover angesiedelt.

daten, das nestor-Siegel für vertrauenswürdige Langzeitarchive (DIN-Norm 31644) und das seit einigen Jahren etablierte DINI-Zertifikat für Open-Access-Repositoryen zu nennen. International streben z. B. EUDAT und RDA einen einheitlichen Rahmen für Zertifikate und Akkreditierungen an.⁵⁹

Die Zahl der zertifizierten oder akkreditierten Dateninfrastrukturen ist bislang überschaubar. Gleichwohl zeichnet sich ab, dass mit Zertifizierungsinitiativen, hierzulande etwa rund um den RatSWD oder in Gestalt der internationalen *Data Seal of Approval*-Community, auch Interessengemeinschaften entstehen, die ihre jeweiligen Standards und Verfahren weiterentwickeln und Erfahrungsaustausch bieten.⁶⁰ Einen niedrigschwelligen Weg, auf eine große Anzahl von Einzelnutzern ausgelegt, bieten Selbstverpflichtungsinitiativen wie jüngst die *Joint Declaration of Data Citation Principles*, welche auf der operativen Seite durch die *Data-FAIRport*-Initiative zur Entwicklung von Tools und Werkzeugen flankiert werden soll.⁶¹ Der Erfolg beider Ansätze ist zum gegenwärtigen Zeitpunkt noch offen. Auch international herrscht im Bereich der Best-Practice-Initiativen Zersplitterung vor (vgl. 2.4).

Dass es zwischen Diensten Synergiepotenziale gibt, steht außer Frage, besonders wenn man die Anwenderschicht der für den Nutzer zugänglichen Dienste und die Datenhaltung, d. h. die physische Speicherung und damit verbundene Dienste, als separierbare Aufgaben betrachtet. Hochschulübergreifende ‚eigene‘ Cloud-Initiativen für die Wissenschaft in Deutschland, der Fusionsprozess im Bereich der Katalogverbände und regionale bzw. fachliche Kooperationen zwischen Hochschulen oder zwischen Hochschulen und außeruniversitären Forschungseinrichtungen bilden das ab. Beispiele sind der Betrieb der GWDG durch die Universität Göttingen und die Max-Planck-Gesellschaft oder der Aufbau des interdisziplinären RADAR-Repositoryums durch ein Konsortium von Universitäten, Helmholtz- und Leibniz-Instituten sowie Informationseinrichtungen.⁶² Auch Cluster-Projekte der EU-Infrastrukturförderung und Ar-

⁵⁹ RatSWD – <http://www.ratswd.de/forschungsdaten/info>; nestor-Siegel – http://www.langzeitarchivierung.de/Subsites/nestor/DE/nestor-Siegel/siegel_node.html; DINI-Zertifikat – <https://dini.de/dini-zertifikat> (zuletzt geprüft am 25.04.2016). RatSWD – <http://www.ratswd.de/forschungsdaten/info>; nestor-Siegel – http://www.langzeitarchivierung.de/Subsites/nestor/DE/nestor-Siegel/siegel_node.html; DINI-Zertifikat – <https://dini.de/dini-zertifikat> (zuletzt geprüft am 25.04.2016)

⁶⁰ Beim RatSWD 31 Forschungsdaten- und Datenservicezentren, bei der Data Seal of Approval-Initiative 57 Dienste (davon ein Drittel deutsche). Vgl. <http://www.ratswd.de/forschungsdaten/fdz> und <http://datasealofapproval.org/en/community> (zuletzt geprüft am 25.04.2016). Alle Zahlen Stand April 2016.

⁶¹ Data Citation Synthesis Group (2014) – Joint Declaration of Data Citation (Webseite). Unterzeichner verpflichten sich nach dem Vorbild der Open-Access-Erklärungen auf die Einhaltung bestimmter Leitlinien für die faire Nachnutzung von Forschungsdaten; Data FAIRport – <http://www.datafairport.org> (zuletzt geprüft am 25.04.2016).

⁶² GWDG – <https://www.gwdg.de>; RADAR – <https://www.radar-projekt.org> (zuletzt geprüft am 25.04.2016). GWDG – <https://www.gwdg.de>; RADAR – <https://www.radar-projekt.org> (zuletzt geprüft am 25.04.2016)

beitsgruppen der Research Data Alliance haben ausdrücklich die Identifikation möglicher übergreifender digitaler Architekturen und Dienste zum Ziel.⁶³

Synergiepotenziale bestehen darüber hinaus zwischen der Forschung und anderen Sektoren, die ähnliche Transformationsprozesse durchlaufen – in Gedächtniseinrichtungen und Archiven oder in der (zumeist wenig vernetzten) öffentlichen Verwaltung. Neben einer technischen und prozessualen Standardisierung, die vernetzten Datenaustausch und Qualitätssicherung ermöglicht, ist v. a. die Frage der Langzeitarchivierung als gemeinsame Aufgabe zu nennen, bei welcher jeweils genau zu prüfende Synergiepotenziale offenkundig sind.

⁶³ Für eine Liste übergreifender Services siehe Field et al. (2013) – Potential of research data.

3 POSITIONEN ZU GRUNDFRAGEN

Wenn der RfII für den komplexen Bereich des Forschungsdatenmanagements auf der Systemebene Empfehlungen ausspricht, so positioniert er sich in offenkundigen, nicht auflösbaren Spannungsfeldern. In der gebotenen Kürze sollen diese Positionen kurz beleuchtet werden – als Antworten auf sich stellende Grundfragen in den jeweiligen Feldern. Diese Antworten sind zugleich Maximen, sie bilden den Hintergrund für die sich anschließenden Empfehlungen des RfII (Kapitel 4).

Top down
vs. bottom up

Dass im Spannungsfeld von *Bottom-up- und Top-down-Ansätzen politischer Steuerung* der Vorrang bei Selbstorganisationsprozessen der Wissenschaft liegen sollte, hat die GWK in ihren Erwartungen an die Arbeit des RfII bereits formuliert. Die Mitglieder des RfII unterstreichen dies im Grundsatz. Angesichts der Herausforderungen des Umgangs mit Forschungsdaten werden jedoch auch Top-down-Impulse gebraucht. Allerdings geht es darum, Prozesse anzustoßen, ohne diese einem zentralistischen Lösungsansatz zu unterwerfen. Ziel ist eine *dynamische Integration verteilten Wissens*,⁶⁴ um gute Wissenschaft zu ermöglichen. Demgemäß schlägt der RfII u. a. die Etablierung koordinierender Akteure und Kooperationen fördernder Anreizsysteme vor. Diese sollen dezidiert in nationalem Rahmen wirksam werden, aber eben unterschiedliche, auch verteilte Akteure und institutionelle Komponenten miteinander verbinden.

Projektförderung
vs. Institutionalisierung

Die klassische förderpolitische Spannung zwischen *Projektförderung und dem Bedarf nach Institutionalisierung* dessen, was systemisch unverzichtbar ist, prägt sich im Bereich des Forschungsdatenmanagements in besonders deutlicher Weise aus. Der Wert digitaler Daten und die Mehrwerte von Datennutzung hängen unmittelbar von der stabil organisierten, zeitlich und technisch dauerhaften Verfügbarkeit der Daten ab. Der RfII befürwortet das Erproben neuer digitaler Verfahren und Methoden auch für den Infrastrukturbereich in Projektform. Gleichwohl kann auf der Basis von Projektförderung kein zeitstabiles, hinreichend durchlässiges und auch ressourceneffizientes Gefüge von Diensten für die Aufbewahrung und die Nutzung von Forschungsdaten in Deutschland entstehen. Der RfII sieht vor diesem Hintergrund alle Beteiligten in der Verantwortung, dort, wo sich Dienste bewährt haben und Kompetenzen vorhanden sind, für dauerhafte Lösungen des Forschungsdatenmanagements deutlich mehr zu tun.

Vereinheitlichung
vs. Vielfalt

Möglicherweise die größte Herausforderung ist in der inneren Heterogenität von disziplinären Anforderungen, Forschungsformen und Methodiken – in der Wissenschaft selbst also – verwurzelt. Sie betrifft das Spannungsfeld zwischen technisch und organisatorisch zwingender *Vereinheitlichung* und einer ebenfalls technisch wie organisatorisch, aber eben auch inhaltlich gebotenen *Vielfalt*, von der die inzwischen über alle ihre Binnendifferenzierungen hinweg di-

⁶⁴ Vgl. Begriffsklärung im Anhang A.

gitalisierte Forschung lebt. Für entscheidend hält es der RfII hier, dass durchgängig Bedingungen für Interoperabilität geschaffen werden. Übergänge bzw. (von den Forschungsinhalten her betrachtet) „Übersetzbarkeiten“, für die wiederum abgestimmte Standards und Schnittstellen die Voraussetzung bilden, stellen das qualitative Merkmal eines Systems dar, das einerseits Vielfalt sichert, andererseits aber variabel – und wachstumsfreudig, Neues ermöglichend („enabling“) – Verknüpfungen eröffnet.

Jenseits technischer Anforderungen gilt Ähnliches, auch wenn sich im Spannungsfeld zwischen *Individuallösungen* und *übergreifenden Diensten* Grundlegendes sinnvoll über Disziplinengrenzen bzw. Forschungsformen hinweg organisieren lässt. Der RfII hält hier den Grundsatz einer standort- und institutionenübergreifenden Denkweise, orientiert an ggf. arbeitsteiligen Diensten, eine deutlich gesteigerte Kooperationsbereitschaft sowie neue Kooperationsformen zwischen wissenschaftlichen Akteuren für sinnvoll. Das Gebot zu kooperieren gilt für Hochschulen untereinander ebenso wie zwischen Hochschulen und außeruniversitärer Forschung, zwischen Wissenschaft, Kultur und öffentlicher Verwaltung oder auch über Regionen und Bundesländer hinweg.

Individuallösungen
vs. übergreifende
Dienste

Das Ziel einer *Offenheit der Datennutzung* schafft ebenfalls ein Spannungsfeld. Denn teils wird man aus guten Gründen *Souveränität über Forschungsdaten* dennoch wollen, etwa zur Qualitätssicherung oder um keine bzw. lediglich gezielte wirtschaftliche Verwertung zuzulassen. Darüber hinaus ist den Belangen *gesetzlichen Datenschutzes Rechnung zu tragen*, und zwar im vollen Umfang deutscher und europäischer Standards. Der RfII bekennt sich im Grundsatz zum Open-Science-Paradigma, wie derzeit zahlreiche Wissenschaftsorganisationen dies tun. Dabei verkennt er nicht, dass der Umgang mit der Systemgrenze zwischen Wissenschaft und Wirtschaft eine Reihe von Fragen aufwirft. Gerade um unter den Akteuren innerhalb der Wissenschaft eine möglichst *offene* Kultur des Teilens von Forschungsdaten zu etablieren, müssen rechtliche Schranken errichtet werden, die auch global verhindern, dass Akteure einen in der Wissenschaft niedrigschwiligen Zugang zu Daten nutzen, um Eigentum an dem zu begründen, was *frei* (im Sinne eines wissenschaftlichen Gemeinschaftsgutes) unter Forschenden zirkulieren soll.

Offenheit vs.
Datensouveränität und
Datenschutz

Zwischen *staatlicher Förderung* und *privatwirtschaftlichem Engagement* gilt es entsprechend mit Weitsicht abzuwägen. Kooperationspartner wie Verlage oder Telekommunikationsunternehmen gehören klassisch zum Wissenschaftssystem hinzu; für Anbieter anderer kostenpflichtiger Mediendienste oder auch von Software gilt längst ähnliches. Freilich können, wo der Staat sich aus Infrastrukturbelangen zu weit zurückzieht, Abhängigkeiten entstehen. Dies ist nicht nur finanziell kritisch, sondern kann auch die Robustheit der Prozesse und die Methodenvielfalt der Forschung selbst bedrohen – etwa durch Monopolbildung, Insolvenzen oder fehlende Redundanz externer Lösungen. Der RfII befürwortet daher, auch Wege der Public Private Partnership zu beschreiten, diese aber

Staatliche Förderung
vs. privatwirtschaftliches
Engagement

umkehrbar zu gestalten. Einem Übermaß an Privatisierung sollte gerade im Bereich dessen, was für Forschungsdaten- und Informationsinfrastrukturen langfristig relevant bleibt, entgegengewirkt werden.

Bund und Länderverantwortlichkeiten

Im Mehrebenen-System der Bundesrepublik Deutschland ergibt sich auch aus der *Aufgabenverteilung zwischen Bund und Ländern* ein Spannungsfeld. Für den Umgang mit ubiquitär anfallenden, kaum aber nur lokal relevanten und *verteilt* keinesfalls zu bewältigenden Mengen heterogener Forschungsdaten sind leistungsfähige, qualitätsgesicherte und effiziente Lösungen nur zu haben, wenn auch die Politik auf Zusammenarbeit setzt. In vielen Fällen würde es den Erfordernissen einer dynamischen Wissensintegration in hohem Maße entgegenkommen, wenn bundesweite Lösungen oder auch Prozesse der Schwerpunktbildung durch intelligente, Grenzen übergreifende Kooperationen der Länder zustande kämen. Probleme des Forschungsdatenmanagements fordern folglich weder stets den Bund noch stets die außeruniversitäre Forschung als dominierenden Akteur. Der Rat sieht Bund und Länder gemeinsam in der Verantwortung, die Herausforderungen im Zusammenhang mit Forschungsdaten und Informationsinfrastrukturen koordiniert anzugehen. Wissensintegration macht nicht an Ländergrenzen halt und erfordert somit gesamtstaatliche Lösungen, die der derzeit vorherrschenden Zerfaserung der Aktivitäten im Bereich des Forschungsdatenmanagements entgegenwirken.

Investition in Infrastruktur vs. Investition in ‚Köpfe‘

Als letztes wichtiges Spannungsfeld ist die finanzierungspolitische *Alternative zwischen der Investition in Infrastruktur und der Investition in ‚Köpfe‘* zu nennen. Im Bereich digitaler Infrastrukturen standen über Jahre Investitionen in Gerätetechnik – Großrechner, Kabelnetze, Software, Lizenzen und Energie – im Vordergrund. Seit einigen Jahren findet hier ein Paradigmenwechsel statt, demzufolge Infrastrukturen an zunehmend komplexen Diensten orientiert, dem Nutzer nahegebracht und finanziert werden müssen: Hardware, Software etc. werden erst im Zusammenspiel mit geschulter, dauerhaft vorgehaltener Expertise zum Werkzeug für die Forschung. In einer zunehmend datenintensiven Wissenschaft müssen neue Tätigkeitsfelder gestaltet werden. Der RfII plädiert daher dafür, den Aufbau von Infrastrukturen für den Umgang mit Forschungsdaten – und zwar auf allen Qualifikationsstufen, welche die Wissenschaft besitzt – stets auch als Sache der Investition in Köpfe anzusehen.

Das schon genannte Zielbild einer dynamischen Integration von Daten, Informationen und Wissen lässt sich nur schrittweise realisieren. So sind Strukturen, Prozesse und Finanzierungspfade in Richtung einer Konvergenz zu entwickeln, die zu Vielgestaltigkeit und auch *Verteiltheit* nicht im Widerspruch steht. Allerdings sind Entschlossenheit und Umsetzungswille gefordert. Denn Stagnation und zunehmend ineffizienter Einsatz öffentlicher Mittel sind die Alternative. Also gilt es, die Schritte baldmöglichst zu gehen.

4 EMPFEHLUNGEN

4.1 VERSTETIGUNG VON INFRASTRUKTURPROJEKTEN

Es entspricht der Entwicklungsdynamik von Digitaltechnologien und der sich unter ihrem Druck verändernden Forschungspraxis, Infrastrukturlösungen forschungsnah und in Projektform zu erproben. Darüber hinaus muss aber deutlich werden, wie Verstetigungsprozesse aussehen können. Ziel sollte sein, ein System nachhaltiger Informationsinfrastrukturen zu schaffen, die sowohl verlässliche Arbeitsstrukturen bieten als auch eine dynamische Weiterentwicklung ermöglichen und eine qualitätsgesicherte Grundversorgung in Deutschland sichern (vgl. auch 4.2).

Von Projektförderungen über längere Zeit abhängige Dienste stellen aus Sicht des RfII ein Risiko für die nachhaltige Infrastrukturgestaltung dar.⁶⁵ Einerseits sollte angestrebt werden, Projektförderung so zu gestalten, dass rechtzeitig vor Ablauf von Projekten Klarheit darüber besteht, ob die Infrastruktur verstetigt wird und in welchen Schritten diese Verstetigung ggf. erfolgt. Die mit der projektförmigen Finanzierung verbundene Kleinteiligkeit von Ansätzen sowie die hohe, mit Kompetenzverlusten verbundene Personalfuktuation stehen dem Bedürfnis nach Konsolidierung, Standardisierung und nachhaltig abrufbarer Expertise entgegen. Wenn es um die dauerhafte Archivierung/Veröffentlichung der eigenen Forschungsdaten geht, fehlt Nutzern im Falle einer ungeklärten Verstetigungsperspektive das Vertrauen. Andererseits kann eine zu langfristige Fixierung hoher Fördersummen nachteilig für einen Wissenschaftsbereich werden, in dem schnell Veränderungen eintreten. Bei uneffektiven Fördermaßnahmen müssen Revisionsmechanismen gegeben sein.

Selbst erfolgreiche Infrastrukturen können sich durch projektförmig eingeworbene Drittmittel allenfalls über Zeiträume von etwa zehn Jahren weiterentwickeln und ‚am Markt‘ halten (vgl. auch 2.3). Daher haben inzwischen wichtige, als Projekt mit Erfolg ‚angeförderte‘ Infrastrukturen wie beispielsweise DARIAH-DE eine strukturelle Grenze erreicht. Eine Trägerschaft für den nachhaltigen Betrieb solcher Forschungsinfrastrukturen ohne oder mit nur schwacher Anbindung an institutionell finanzierte Einrichtungen scheint unter den aktuellen Rahmenbedingungen in Deutschland leider nicht realisierbar. Es fehlt an einer geeigneten Trägerstruktur.

⁶⁵ Diese Einschätzung basiert u. a. auf Ergebnissen des DFG-Workshops „Erfolgskonzepte wissenschaftlich relevanter Informationsinfrastrukturen“ (2014).

4.1.1 Um den Einsatz öffentlicher Mittel in Infrastrukturprojekte sinnvoll zu flankieren, empfiehlt der RfII ein Phasenmodell, das Pfadentscheidungen für die mögliche Überführung in eine geeignete Trägerschaft bzw. die Konsolidierung ermöglicht. Der Übergang aus der Projektphase in die Verstetigung/Konsolidierung muss dabei evaluationsbasiert erfolgen. Ziel der Evaluation ist es zunächst, die wissenschaftliche Qualität, Akzeptanz und Relevanz der fraglichen Infrastruktur zu betrachten. Generell hält der RfII die Einhaltung der FAIR-Data-Prinzipien für eine sinnvolle Richtschnur.⁶⁶ In einem zweiten Schritt sind einzelfallbasiert Verstetigungsoptionen zu prüfen. Hierbei sind durch geeignete Gutachter Empfehlungen für den weiteren Betrieb und eine dauerhafte Konsolidierung einschließlich der personellen Ressourcen zu geben (vgl. 4.1.2 und 4.1.3). Ebenso sind wissenschaftspolitische Gesichtspunkte zu betrachten: fachliche, geografische, institutionelle Schwerpunktbildung einerseits sowie andererseits Sicherung von hinreichender Pluralität und forschungsinterner Konkurrenz, dazu das Entwicklungspotenzial der Projekte in der neuen Struktur. Dieses Phasenmodell sollte auch dann leitend sein, wenn die Projektphase aus institutionellen Mitteln einer außeruniversitären Forschungseinrichtung finanziert wird.

Auch wenn Forschungsprojekte idealerweise stets so aufgesetzt und durchgeführt werden sollten, dass die dabei entstehende Forschungsdatensammlung in eine übergreifende Forschungsdateninfrastruktur übergeleitet oder integriert werden kann, ergeben sich doch im Hinblick auf eine Verstetigung mindestens folgende Optionen:

- Weiterführung als singuläre Infrastruktur;
- Angliederung an eine bestehende Infrastruktur;
- Integration in einen zentralen, im Rahmen einer nationalen Infrastruktur bereits bestehenden Dienst zur langfristigen Nachnutzung (vgl. 4.2);
- Zusammenfassung unterkritisch ausgestatteter Infrastrukturen in übergreifenden Betriebsmodellen;
- Archivierung des erreichten Stands;
- Verbesserung der Kosten-Nutzen-Relation durch Rückbau auf einen auch in einer längerfristigen Perspektive prioritären Kern.

Zur besseren Planbarkeit dieser modellhaften Alternativen müssen für alle Projekte übergreifende Mindeststandards zur Datenhaltung entlang des Datenlebenszyklus gelten. Weitergehende Anforderungen, die in einem Projekt entstehen, müssen mittelfristig in den übergreifenden Standard integriert werden. So

⁶⁶ FAIR = Findable, Accessible, Interoperable, Reusable: <https://www.force11.org/group/fair-group/fairprinciples> (zuletzt geprüft am 25.04.2016). Vgl. auch Darstellung der Data-FAIRport-Initiative in Abschnitt 2.5.

lassen sich Prozesse der Standardisierung und Normierung während des Aufbaus und Betriebs der jeweiligen Dateninfrastrukturen aktiv unterstützen.

Um einerseits kritische Finanzierungslücken zu vermeiden und andererseits keine falschen Anreize durch variable Projektförderbedingungen zu schaffen, muss bei erfolgreichen Projekten frühzeitig ein Übergang in geeignete Trägerschaft oder in die institutionelle Förderung, nämlich in die Grundfinanzierung der tragenden Einrichtung, organisiert werden.

4.1.2 Im schematischen Werdegang kann der Aufbau einer Infrastruktur aus institutioneller Förderung oder über konsekutive Drittmittelförderungen gleichermaßen erfolgen. Die Begutachtung in der Projektphase ist in der Regel Sache des jeweiligen Mittelgebers. Sollbruchstellen entstehen durch Zuständigkeitslücken, z. B. wenn kein geeignetes Förderprogramm vorhanden ist und nicht ein, sondern mehrere Förderer bzw. Träger Finanzierungsanteile übernehmen sollen. Wird Institutionalisierung erwogen, muss eine unabhängige Instanz die Pfadentscheidung über die nächste Betriebsphase durch eine Evaluation begleiten. Es wird empfohlen, solche Evaluationen auch für kleinere Infrastrukturprojekte durchzuführen. Hiermit können geeignete Akteure im Wissenschaftssystem beauftragt werden, die neben wissenschaftlicher Exzellenz auch die Bedarfe des Wissenschaftssystems insgesamt sowie weitere relevante Gesichtspunkte in den Blick zu nehmen haben (vgl. 4.1.3).

Unabhängige
Begutachtung

4.1.3 Das System der wissenschaftlichen Begutachtung in Deutschland ist traditionell auf thematisch und methodisch *disziplinäre* und auf *Forschungsinhalte* angelegte Evaluationsvorgänge ausgerichtet. Für die Bewertung von Informationsinfrastrukturen gibt es zwar erste, aber nur wenig erprobte oder allgemein anerkannte Konzepte. Gleichmaßen kann bisher nicht ohne weiteres auf Gutachterpools mit einschlägiger disziplinenübergreifenden Expertise zurückgegriffen werden.

Konzept für
regelmäßige
Begutachtungen

Der Rfll empfiehlt daher die Entwicklung eines Konzepts, das für regelmäßige Begutachtungsvorgänge von Forschungsdateninfrastrukturen die jeweiligen fachmethodischen und fachkulturellen Belange berücksichtigt, aber grundlegende Ziele und Kriterien für eine infrastrukturbezogene Evaluation benennt. Neben der wissenschaftlichen Qualität und der Nutzerorientierung sollten auch gesellschaftliche Relevanz und das wirtschaftliche Nutzungspotenzial der Datenbestände bewertet werden.

4.2 SCHWERPUNKTBILDUNG UND ARBEITSTEILUNG IM SYSTEM

Die aktuelle Situation für den Umgang mit Forschungsdaten in Deutschland ist von vielen, umständehalber vielfach schwach koordinierten Aktivitäten und

Investitionen geprägt. Die Nachfrage nach koordinierten Services steigt. Noch sind nicht genug Dienste aufgebaut, um flächendeckend ein verbindliches Forschungsdatenmanagement etablieren zu können. Soll bundesweit ein effizienteres System entstehen, sind Umbauten und eine Aggregation funktionaler Art erforderlich (vgl. 4.1.1).

Eine solche Aggregation kann wegen der hohen Komplexität nicht über einen Masterplan entworfen und *top down* umgesetzt werden. Um neue, arbeitsteilig organisierte Strukturen zu schaffen, bedarf es vielmehr einer Kombination von starken Vernetzungsanreizen und einer zumindest indirekt wirksamen Koordination. Mittels eines übergreifenden Prozesses gilt es vor allem, unterstützende Strukturen zu entwickeln sowie vorhandene Expertise in Bezug auf Datenkuratierung und Analysemethoden zu bündeln.

Ziel ist es, die Kapazitäten für das Forschungsdatenmanagement auszubauen und langfristig zu einem integrierten System von Forschungsdateninfrastrukturen zu kommen. Der RfII skizziert hierzu den Umriss, ohne Umsetzungsentscheidungen im Detail vorwegzunehmen.

Grundversorgung mit Speicherinfrastrukturen und Services

4.2.1 Viele Aspekte des Forschungsdatenmanagements sind generischer und damit übertragbarer Natur, auch wenn die im Wissenschaftssystem generierten Daten selbst insgesamt sehr heterogen sind. Mit Blick auf Kosten und Effizienz können und sollten generische Dienste arbeitsteilig aufgebaut und angeboten werden. Der RfII schlägt daher die Gründung eines Verbundes vor, der die vorhandenen Kompetenzen bündelt und für eine Grundversorgung mit Speicherinfrastrukturen und Services sowie einen schnellen Kompetenztransfer im Wissenschaftssystem sorgt. Diese *Nationale Forschungsdateninfrastruktur* (NFDI) sollte die Form eines Netzwerkes haben, disziplinen- bzw. communityübergreifend⁶⁷ angelegt sein und sowohl die existierenden großen Informationseinrichtungen als auch die nationale Ebene der ESFRI-Projekte und die Repositorien weiterer, in ihren Bedarfen hinreichend homogener Nutzergruppen einbinden. Als Verbund muss die NFDI eine auf Handlungsfähigkeit und die Steuerung von Entwicklungen (hier zunächst: eines Transitionsprozesses) ausgelegte Governance besitzen. Gleichwohl sind Partizipationsmöglichkeiten vorzusehen, die der Einbindung einer Vielfalt von Akteuren angemessen Rechnung tragen (vgl. 4.4).

Nationale Forschungsdateninfrastruktur

Fragmentierung überwinden

4.2.2 Die *Nationale Forschungsdateninfrastruktur* (NFDI) ist eine vernetzte, verteilte Infrastruktur für Datenerzeuger und Nutzer, fachliche Services sowie Datendienste mit einem gemeinsamen Nutzerzugang und einer strategisch aufgestellten, entscheidungsfähigen Koordinationseinheit. Eine offene, auf Integ-

⁶⁷ Vgl. die Begriffsklärung im Anhang A: Der engeren, eher sozial gefassten Einheit der (Fach-)Community steht der weitere, eher taxonomische Begriff der Disziplin gegenüber.

ration angelegte, *evolviende* Struktur erlaubt die Kombination von Daten aus unterschiedlichen Communities für weitergehende wissenschaftliche Analysen, einschließlich Big-Data-Anwendungen, und ist darauf ausgelegt, Fragmentierung zu überwinden. Das Grundprinzip eines arbeitsteiligen Vorgehens ermöglicht sinnvolle Spezialisierung, Schwerpunktbildung und die Gründung von Zentren bzw. „Knoten“ zur Wahrnehmung übergeordneter Aufgaben.

Als funktionaler Verbund hat die NFDI folgende nationale, das Gesamtsystem betreffende Herausforderungen anzugehen:

- Die Definition von übergreifenden Mindeststandards und von Qualitätsmanagement bei der Datenbeschreibung und -haltung;
- die Entwicklung von generischen Datenanalyseverfahren und ihre Übertragung auf spezifische Daten und Fragestellungen;
- Entwicklung, Aufbau und Bereitstellung von generischen Datenservices und Datenspeichern sowie Schnittstellen zur Anbindung verteilter lokaler Repositorien;
- Aus-, Fort- und Weiterbildung.⁶⁸

Zu berücksichtigen ist die Anschlussfähigkeit an europäische Entwicklungen, namentlich die European Open Science Cloud, sowie die Verbindung zu bereits bestehenden Repositorien und Strukturen und zu einzelnen disziplinären und spezifischen Diensten. Hierzu sind Schnittstellen erforderlich, die die verschiedenen Datenschichten verbinden und eine Anbindung von generischen Diensten für das Management der Forschungsdaten ermöglichen. Neben der technischen Infrastruktur zur Datenspeicherung und der professionellen Nutzerunterstützung können auch Methoden und Software zum Umgang mit Daten und für die Datenanalyse kollaborativ entwickelt werden (Recherche, Aggregation, Visualisierung etc.).

Es wird empfohlen, beim Aufbau der NFDI die technische Ausstattung, den Support und die Methodenentwicklung/Datenanalyse im Zusammenhang zu sehen und zu planen. Hier kann auf Konzepte zum Aufbau des Nationalen Hoch- und Höchstleistungsrechnens (NHR) zurückgegriffen werden.⁶⁹

IT, Methodenentwicklung und Service im Zusammenhang planen

Aufgrund der hohen Dynamik der Entwicklungen und der letztlich unklaren Anforderungen, die zukünftig an das Forschungsdatenmanagement gestellt werden, sind die Prozesse zur weiteren Entwicklung der NFDI mit der notwendigen Offenheit zu gestalten. Gleichzeitig gilt es, diese Infrastruktur mit einem breiten Kreis der Wissenschaft abzustimmen, um vorhandene Strukturen und die Bedarfe soweit wie möglich einbeziehen zu können.

⁶⁸ Aufzählung ohne Anspruch auf Vollständigkeit, z. B. kommen mit Datenschutz und Datensicherheit verbundene Aufgaben – wie im Forschungsdatenmanagement durchgehend wichtig – hinzu.

⁶⁹ WR (2015) – Empfehlungen zur Finanzierung des NHR.

4.2.3 Bund und Länder sowie ggf. Kommunen und private Stiftungen als Träger wissenschaftlicher (oder wissenschaftsrelevanter behördlicher) Einrichtungen verfügen über unterschiedliche Möglichkeiten und Voraussetzungen für die Einrichtung von Forschungsdateninfrastrukturen. Angesichts des großen institutionellen Bedarfs (vgl. 2.3) wird ihnen empfohlen, gemeinsam mit Datenproduzenten, Betreibern und Datennutzern übergreifend arbeitende Infrastrukturzentren zu entwickeln, die eine Grundversorgung an Forschungsdatenrepositorien mit entsprechender Nutzerunterstützung sicherstellen. Diese Zentren sollen zur Kopplung mit internationalen Initiativen befähigt sein. Gerade im derzeit entwicklungsstarken Bereich der Forschungsdatenrepositorien, in denen Forschungsdaten aus Einzelprojekten *lege artis* hinterlegt werden sollen, haben lokale Strukturen nur durch die Anbindung an übergreifende Repositorien langfristig einen wissenschaftlichen Mehrwert. Aufgaben wie die Langzeitarchivierung (siehe 4.3) sowie der Betrieb von großen Forschungsdatensammlungen benötigen Kontinuität sowie ggf. eine kritische Menge an Expertise und Ressourcen. Entscheidern auf der Ebene der Leitung von Hochschulen, Forschungs- und Infrastruktureinrichtungen wird empfohlen, ihrerseits ebenfalls möglichst in (ggf. zu schaffenden) Verbundstrukturen zu arbeiten, die in die NFDI eingebettet werden können bzw. die Anbindung an größere Infrastrukturzentren zu suchen. Insbesondere Letzteres gilt auch für die Entwicklung institutionsübergreifender Dateninfrastrukturen in Fachgemeinschaften und Forschungsverbänden.

4.2.4 Die hohe Spezifik der Aufgaben wird einen schrittweisen Aufbau der NFDI erfordern. Im ersten Schritt werden übergreifende Standards⁷⁰ der Datenbeschreibung gebraucht, die außerhalb einer Community oder auch über die Grenzen von Forschungsformen hinweg verstanden werden und Forschenden die notwendigen Informationen geben, um zu beurteilen, ob die Daten für ihre Forschung von Interesse sind (vgl. dazu auch 4.6.1).

Soweit Mindestanforderungen erfüllt sind, können weitere Schritte der technischen Vereinheitlichung von an die NFDI anzubindenden Infrastrukturen gegangen werden. Damit wird einem breiten Kreis von Forschenden der Zugang zu verschiedenen Ressourcen ermöglicht und der Grundstein für die Kombination und damit Auswertung von Daten aus unterschiedlichen Quellen gelegt. Zu den Mindestanforderungen gehören gemeinsame Standards für die Qualität der Daten sowie deren Beschreibung mit Metadaten; dokumentierte Regelungen für Zugriffsrechte (ggf. im Sinne von Open Data) und für die Zweitverwendung.

⁷⁰ Zum Verständnis von Standards vgl. auch die Begriffsklärung im Anhang A.

Es wird empfohlen, für die angebotenen Dienste universellen Zugang für registrierte Anwender zu schaffen, zentrale Dienste für Zugriffs- und Schreibrechte zu etablieren und den Datenzugriff für Nutzer verschiedenen Typs übergreifend zu regeln. Weitere Ermöglichungsstrukturen sind frühzeitig über ein Fachkonzept zu definieren (für Beispiele siehe Anhang D.3). Es wird empfohlen, Erfahrungen aus der Etablierung ähnlich gelagerter Initiativen aufzuarbeiten.⁷¹

Gemeinsam
betriebene
Dienste

4.2.5 Größere Infrastrukturzentren können als Dienstleistungen für den Aufbau von Repositorien sowohl Hardware als auch methodisches Know-how und Beratung bereitstellen. So sind Skalierungseffekte zu erwarten, die im Verbund mit der Professionalisierung der Dienste zu mehr Kosteneffizienz führen und die Datenqualität sichern. Der RfII empfiehlt, dass die Infrastrukturanbieter durch Methodenkompetenz und überzeugenden Service Anreize schaffen, die Forschende vom Mehrwert der angebotenen Dienstleistung überzeugen. Dass guter Umgang mit Forschungsdaten als selbstverständlicher Teil guter wissenschaftlicher Praxis gelebt wird, kommt hinzu. Die Hoheit über die Daten muss dabei immer bei den Daten erzeugenden Forschenden bleiben. Wo Forschende *bottom up* neue Repositorien aufbauen, wird ihnen empfohlen, kostenbewusst zu agieren und existierende Services von Infrastruktureinrichtungen zu nutzen.

Dienstleistungen für
gute wissenschaftliche
Praxis

4.2.6 Innerhalb der Forschungsdateninfrastrukturen wird es nicht nur Arbeitsteiligkeit, sondern eine funktionale Differenzierung geben. Hier bietet sich an, hinsichtlich der Komplexität der Datenhaltungsaufgabe, der fachlichen Spezialisierung und des Umfangs der Nutzerunterstützung/dynamischen Wissensintegration Unterscheidungen zu treffen. Der RfII hält folgende, nicht abschließende Differenzierung für sinnvoll:

Föderierte Struktur
und Anbindung lokaler
Dienste

- Auf die breit aufgestellten Infrastrukturzentren für die NFDI-Grundversorgung wurde bereits unter 4.2.4 eingegangen. Es wird empfohlen, wettbewerbliche Steuerungselemente zu schaffen, damit diese Infrastrukturzentren um Nutzer und Mittel konkurrieren. Dadurch können sich Kompetenzzentren mit unterschiedlichen Schwerpunkten herausdifferenzieren, die die speziellen Anforderungen großer Communities berücksichtigen und gleichzeitig mit ihren generischen Angeboten den Bedarf kleinerer Communities abdecken. Solche Kompetenzzentren können z. B. in Einrichtungen mit bereits bestehenden großen Community-Repositories entstehen.⁷²

⁷¹ Z.B. Lösungsansätze aus Vorhaben wie OpenAIRE (<https://www.openaire.eu>), Europeana (<http://www.europeana.eu>) und Deutsche Digitale Bibliothek (<http://www.deutsche-digitale-bibliothek.de>, alle Links zuletzt geprüft am 25.04.2016).

⁷² Der Übergang zwischen (eher als breite Dienstleister aufgestellten) Infrastrukturzentren und (primär auf spezifische Bedarfe reagierenden) Kompetenzzentren ist nicht trennscharf, sondern kann im Einzelfall durchaus fließend sein.

- In Teilen des Wissenschaftssystems entstehen großvolumige Datensammlungen (Hochdurchsatzmessungen, Simulationen). Es wird empfohlen, Datenzentren mit hohen Analyse- und Verarbeitungsaufwänden perspektivisch mit entsprechend ausgestatteten HPC-Kompetenzzentren zu verbinden.
- In einer nationalen, föderierten Infrastruktur wird weiterhin auch eine Grundversorgung mit Diensten durch kleinere Akteure zu finanzieren sein. Hier sind beispielsweise Bibliotheken, Archive, kleinere Rechenzentren oder IT-Infrastrukturen an wissenschaftlichen Instituten zu nennen. Diese bilden eine eigene, ebenfalls auf Kooperation angelegte Ebene in der NFDI und sind durch geeignete Organisationsstrukturen in funktionaler Weise in das System einzubinden.
- Wissenschaftliche Datensammlungen entstehen geografisch verteilt an Einrichtungen unterschiedlichen Typs. Es wird empfohlen, diese Ressourcen perspektivisch an fachspezifisch passende Kompetenzzentren oder eben größere Infrastrukturzentren anzubinden (vgl. 4.2.4). Es ist im Einzelfall zu prüfen, ob Daten für einen effizienteren Zugriff oder zur langfristigen Speicherung in größere Datenzentren physikalisch integriert werden sollten. Grundsätzlich sollte kleineren Forschungsdatenzentren, die z. B. auf Ebene von Arbeitsgruppen betrieben werden, die Integration in geeignete Trägereinrichtungen offenstehen.
- Der Ausbau von Datenspeicherdiensten (wissenschaftlichen Clouds) muss langfristig und bedarfsorientiert erfolgen, um den Anwendern den jeweils größten Nutzen bieten zu können. Es wird empfohlen, kleinere lokale Speicherinfrastrukturen zunächst virtuell zu verbinden, indem vereinheitlichte Datenzugriffsprotokolle implementiert werden. Langfristig sollten auch solche, lokal gehaltenen Daten migriert werden, um eine hinreichend langfristige Datenverfügbarkeit zu gewährleisten und Kosten zu senken.

Die großen Infrastrukturzentren sollten in der Lage sein, füreinander Backup-Funktionen zu übernehmen. Es wird empfohlen, Aufbau, Betrieb und Ausbau der Infrastrukturzentren an regelmäßige Evaluationen zu koppeln.

Langfristiger Transitionsprozess

4.2.7 Die Finanzierung der NFDI ist eine Herausforderung, die den normalen Rahmen von Weiterentwicklungsprozessen für den Forschungsbetrieb übersteigt. Sie ist über eine unverzichtbare, allerdings nur in Grenzen mögliche Steigerung des Eigenengagements der institutionell geförderten wissenschaftlichen Akteure hinaus auch eine gemeinsame Aufgabe von Bund und Ländern. In der Breite des Wissenschaftssystems gilt es, diese Infrastruktur abzustimmen, um vorhandene Strukturen und die Bedarfe soweit wie möglich einbeziehen zu können. Der RfII geht von einem Transitionsprozess aus, der 15-20 Jahre umfasst.

Daher sind für den NFDI-Verbund adaptive und langfristig funktionale Koordinationsstrukturen sowie geeignete Pilotprojekte zu entwickeln. Aufgrund der hohen Dynamik der Entwicklungen sind die Prozesse zur weiteren Entwicklung der NFDI mit der notwendigen Offenheit zu gestalten. Der RfII empfiehlt, nach einer Konkretisierung von Bedarfen die Aufnahme in die Nationale Roadmap für Forschungsinfrastrukturen zu prüfen (vgl. 4.13).

4.3 LANGZEITARCHIVIERUNG UND LANGZEITVERFÜGBARKEIT VON FORSCHUNGSDATEN

Nicht nur in den Bereichen des kulturellen Erbes und generell des Publizierens, sondern auch für die während des Forschungsprozesses eingesetzten und erzeugten digitalen Daten besteht die Herausforderung, eine langfristige Verfügbarkeit, Überprüfbarkeit und Nutzbarkeit ausgewählter Forschungsdaten⁷³ sicherzustellen. Damit sind komplexe Aufgaben verbunden, die weit über die reine Speicherung hinausgehen und zusätzliche IT-Ressourcen binden.⁷⁴ Zu programmieren, zu betreiben und (weiter) zu entwickeln sind Architekturen, Prozesse und Dienste, die zum einen übergreifende und disziplinunabhängige Funktionen sicherstellen (Qualitätssicherung, Standardisierung nach international gültigen Kriterien, Datenpersistenz, Zugriffssicherheit im Sinne von Authentifizierung/Autorisierung, Suche). Zum anderen müssen disziplinabhängige Aspekte berücksichtigt werden, wie die Dokumentation des fachspezifischen Kontextes der Daten.⁷⁵ Das Kompetenznetzwerk nestor hat in seinem Positionspapier 2015⁷⁶ auf einer allgemeinen Ebene zentrale Aufgaben des Handlungsfeldes Langzeitarchivierung für die digitale Forschung und Bildung benannt. Die Sicherstellung der Langzeitverfügbarkeit speziell von Forschungsdaten ist mit Anforderungen verbunden, die darüber noch hinausgehen.

4.3.1 Der RfII regt einen Fachdiskurs über die Differenzierung zwischen projektlaufzeitnaher Speicherung sowie über deutlich länger zugeschnittene Archivierungsfristen an. Die gebotenen großen Zeitspannen der Archivierung digitaler Forschungsdaten sind derzeit bei Weitem nicht ausreichend berücksichtigt. Die DFG nennt in ihren Empfehlungen zur „Sicherung guter wissenschaftlicher Praxis“ das Vorhalten von Primärdaten über einen Zeitraum von zehn Jahren (nicht zwangsläufig bei offenem Zugang).⁷⁷ Diese Aufbewahrungsfrist sichert die Nachvollziehbarkeit von Forschungsprozessen und die Überprüfbarkeit von Er-

Fachdiskurs über Archivierungsfristen

⁷³ Auswahl heißt: Hier stellt sich die Frage nach Relevanzkriterien (siehe dazu knapp 4.3.3).

⁷⁴ Einen Eindruck der vielfältigen Anforderungen vermitteln die im Anhang C dargestellten Szenarien.

⁷⁵ Neuroth et al. (Hg.) (2012) – Langzeitarchivierung.

⁷⁶ Nestor (2015) – Positionspapier Digitale Langzeitarchivierung.

⁷⁷ DFG (2013) – Sicherung guter wissenschaftlicher Praxis, S. 21.

gebnissen – aber nur relativ zeitnah. Für viele Disziplinen sowie grundlegende Forschungsprobleme ist der Zeitraum deutlich länger, in dem Forschungsdaten für neue Auswertungen oder auch zur Überprüfung von Ergebnissen verfügbar sein sollten.

4.3.2 Grundsätzlich sind bei der digitalen Langzeitarchivierung vielfältige Aspekte zu beachten und Leitlinien gefordert. Zunächst gilt es, den Erhalt der Daten (so wie sie gespeichert wurden) gegen Risiken zu sichern. Verfahren hierzu sind eine redundante – ggf. auch geografisch verteilte – Datenspeicherung sowie Verfahren zur regelmäßigen Überprüfung und ggf. Wiederherstellung der Integrität der Daten. Bei überschaubaren Zeiträumen von einigen Jahren reichen solche Maßnahmen aus, wobei früher oder später Datenmigrationen bei Wechsel der physischen Datenträger hinzukommen. Je länger eine Aufbewahrung erfolgen soll, desto wahrscheinlicher sind Maßnahmen erforderlich, welche die Daten nicht nur technisch, sondern auch inhaltlich nutzbar erhalten (sogenannte *Content Preservation*). Notwendige Maßnahmen können z. B. Migration der Daten in neue Formate oder Anpassungen an neue Hard- und Softwareumgebungen sein. Werden Datensätze in Formaten abgelegt, die in neuen Softwareumgebungen nicht mehr lesbar sind, ist eine langfristige Nutzung allenfalls mittels aufwändiger Umformatierungen (Formatmigration) möglich. Um die Häufigkeit solcher Formatmigrationen zu minimieren, erfolgt die Archivierung von Datenobjekten in langfristig verarbeitbaren Formaten (z. B. csv, xml). Angesichts der hohen Vielfalt der in den verschiedenen Disziplinen verwendeten Metadaten-Formate hält der RfII in diesem Feld eine stärkere Einigung auf Standards für dringlich. International ist auf diesem Gebiet z. B. die Research Data Alliance (RDA) aktiv.

Forschenden und ihren Einrichtungen wird empfohlen, in Kooperation mit den Dienstleistern der NFDI frühzeitig eine entsprechende Planung für den Umgang mit Daten und die entlang des Datenlebenszyklus benötigten Werkzeuge, Services und Workflows zu entwickeln (zu Datenmanagementplänen vgl. auch 4.6.2). Datenproduzenten (Forschende) sollten an ihren jeweiligen Standorten künftig Werkzeuge zur Verfügung haben, die sie bei der Einrichtung ihrer Daten auf Belange der Langzeitverfügbarkeit (einschließlich der sogenannten *Kuration*) unterstützen. Solche Werkzeuge zu entwickeln und bereitzustellen, ist Aufgabengebiet von fachspezifischen wie disziplinunabhängigen Datenzentren im Rahmen der NFDI.

4.3.3 Für Entscheidungen über den Umfang bzw. die Volumina der Langzeitarchivierung (ggf. mit Erhaltung und Aufbereitung) und die Wege zur Sicherstellung der Langzeitverfügbarkeit müssen wissenschaftliche, insbesondere fachkulturelle Erwägungen mit unvermeidlichen Wirtschaftlichkeitsüberlegungen verbunden werden. Grundsätzlich ist von der Prämisse auszugehen, dass nicht alle Daten für immer gespeichert werden müssen. Generische Kosten-Nutzen-Abschätzungen für Langzeitarchivierung sind gleichwohl schwierig, da die

mögliche wissenschaftliche Relevanz (aber auch der mögliche wirtschaftliche Wert) von Forschungsdaten in der Zukunft kaum sicher eingeschätzt werden kann. Eine Relevanzabschätzung für Forschungsdatensammlungen können aus Sicht des Rfll in erster Linie die jeweiligen Disziplinen selbst durchführen. Alternativ müssen – insbesondere bei *verwaisten* Daten, vergleichbar dem Umgang mit analogen Beständen – Datenarchivare nach zuvor festgelegten Kriterien über Erhalt oder Vernichtung von Daten und Datensammlungen entscheiden.

4.3.4 Langzeitarchivierung und Sicherung der Langzeitverfügbarkeit gehören zu den wichtigen Aufgabengebieten einer Nationalen Forschungsdateninfrastruktur. Der Rfll empfiehlt, im Bereich der Langzeitverfügbarkeit von Forschungsdaten die Problematik der (nur) projektförmigen Finanzierung von potenziell nachhaltigen Datenerhebungen besonders zügig anzugehen. Für die Langzeitverfügbarkeit erforderliche Ressourcen sollten auf dem Wege einer langfristig abgestimmten Finanzierungsplanung bereitgestellt werden, die ein schrittweises Vorgehen vorsehen kann, aber der nationalen Bedeutung der Aufgabenstellung Rechnung trägt. Zu leisten und zu koordinieren sind Aufgaben der Langzeitarchivierung als informationsinfrastrukturelle Daueraufgaben durch ggf. miteinander vernetzte Einrichtungen im Rahmen der NFDI (vgl. 4.4).

Langzeitarchivierung
als Daueraufgabe
organisieren

Bis auf relativ wenige – i. d. R. disziplinspezifische – Datenzentren existieren in Deutschland noch keine etablierten Rollen und Verantwortlichkeiten, weswegen konkrete Zuständigkeiten geklärt werden müssen. Außerdem ist das Problem zu lösen, welche Aufgaben disziplinbezogen und welche disziplinenübergreifend zu erledigen sind. Einrichtungen, die Langzeitarchivierung für Forschungsdaten anbieten, müssen überdies den Umfang dieser Dienste (*Service Levels*) und die Anforderung an die zu liefernde Datenqualität hinreichend transparent beschreiben. Ein innerhalb der NFDI eng abgestimmter Prozess, der diese Aufgaben erfüllt, sollte Synergien zwischen Wissenschaft und Einrichtungen der Informationsinfrastruktur bei der Etablierung von Services identifizieren und nutzen. Bislang konnten zuständigkeitsbedingte Blockaden im Feld der Sicherstellung von Langzeitverfügbarkeit nicht überwunden werden, daher ist die Schaffung klar definierter und kommunizierter Rollen und Verantwortlichkeiten auf der Basis nachgewiesener Kompetenzen hier besonders dringlich.

4.3.5 Sofern Langzeitarchivierung außerhalb der Einrichtung geleistet wird, welche die Daten erzeugt hat – d. h. Langzeitarchivierung als Dienstleistung erfolgt – werden geeignete Kosten- und Geschäftsmodelle benötigt. Der Rfll hält, um Langzeitverfügbarkeit nachhaltig zu etablieren, institutionelle Anstrengungen sowie eine Unterstützung durch Bund und Länder für unabdingbar, er befürwortet aber eine angemessene Kostenbeteiligung der nutzenden Einrichtungen.

Kostenbeteiligung
nutzender
Einrichtungen

Beitragsmodelle können auf unterschiedliche Speicherfristen und Nutzergruppen abgestimmt sein. Das EU-Beihilferecht ist zu beachten, sofern es um Dienstleistungen geht, die nicht auf Services innerhalb der Wissenschaft beschränkt sind, sondern sich auch an privatwirtschaftliche Akteure richten.

4.4 NEU ZU ETABLIERENDE AKTEURE/ZUSTÄNDIGKEITEN

Für den Aufbau einer netzwerkförmigen Nationalen Forschungsdateninfrastruktur (NFDI) sind grundlegende Verständigungsprozesse und arbeitsteilige Strukturen zu organisieren. Auch wenn die NFDI darauf angelegt ist, bestehende Akteure auf integrativem Wege zusammenzubinden, sind neue Rollen und Zuständigkeiten zu etablieren.

Schrittweise Netzwerkbildung

4.4.1 Die NFDI soll als Netzwerk schrittweise und bedarfsorientiert entstehen. Benötigt wird ein initiales Konstrukt (*Gründungsverbund*) im Sinne einer Kerngruppe. Ebenso muss von Anfang an Verbindlichkeit hergestellt werden und Ressourcenfragen müssen beantwortet werden können, damit aus Verständigungen operative Umsetzungen folgen. Bund und Länder sollten in ihrer Zuständigkeit als Träger von Forschungseinrichtungen und Hochschulen die Kooperation und die gute Praxis des Forschungsdatenmanagements in den von ihnen getragenen Einrichtungen in geeigneter Weise mandatieren. Der RfII empfiehlt Bund und Ländern ein koordiniertes Vorgehen anhand eines geeigneten Instruments, z. B. einer Roadmap für die NFDI. Geeignete Akteure sollten hierzu zeitnah ein Strategiepapier verfassen. Das Feld der Langzeitarchivierung bedarf besonderer Anstrengungen und eines hohen Maßes an institutioneller Kontinuität, nicht zuletzt um Skalierungs- bzw. Synergieeffekte zu realisieren (vgl. 4.3.4). Der RfII empfiehlt hier die frühzeitige Bündelung von Ressourcen in einem Netzwerk von auf spezifische Bedarfe der Langzeitverfügbarkeit eingestellten Kompetenzzentren bzw. übergreifend arbeitenden Infrastrukturzentren innerhalb der NFDI. Bund und Länder sollten bei der Gründung von Kompetenzzentren für die Langzeitarchivierung zusammenwirken und die Kapazitäten aufeinander abstimmen.

Gemeinsames Vorgehen von Bund und Ländern

Schaffung von Ver- bundstrukturen zur Mitwirkung in der NFDI

4.4.2 Zugunsten einer Mitwirkung in der NFDI verfügen die Träger wissenschaftlicher (oder wissenschaftsrelevanter behördlicher) Einrichtungen über unterschiedliche Möglichkeiten und Voraussetzungen, gemeinsam mit Datenproduzenten, Betreibern und Datennutzern eine Grundversorgung an Forschungsdatenrepositorien zu entwickeln (vgl. 4.2.3). Entscheiden auf der Ebene der Leitung von Hochschulen, Forschungs- und Infrastruktureinrichtungen wird empfohlen, sich in Verbundstrukturen einzubringen, die in die NFDI integriert werden können. Diese Empfehlung gilt auch für die Entwicklung institutionsübergreifender Dateninfrastrukturen in Fachcommunities und Forschungsverbänden.

4.4.3 Evaluation und Zertifizierung von bzw. für Informationsinfrastrukturen sind im Wissenschaftssystem noch nicht in der erforderlichen Breite etabliert (siehe Kapitel 2.3 und 2.5 sowie 4.1.1). Der RfII empfiehlt, dass sich die Wissenschaftsorganisationen und ggf. der NFDI-Gründungsverbund zunächst verständigen, welche übergreifenden Kriterien und Zertifizierungen zur Vernetzung auf nationaler Ebene beitragen können und in ihren jeweiligen Verfahren etabliert werden können. Ebenso schlägt der RfII den zeitnahen Aufbau eines national und international zusammengesetzten, disziplinenübergreifenden Gutachterpools vor (vgl. 4.1).

Verständigung über
Evaluationskriterien

4.4.4 Um vorhandene Strukturen und Bedarfe im Hinblick auf eine nationale Forschungsdateninfrastruktur abzustimmen, ist ein breiter Diskurs notwendig. Der private Sektor ist in geeigneter Weise an den Entwicklungsprozessen für eine Nationale Forschungsdateninfrastruktur zu beteiligen, da den privatwirtschaftlich organisierten Informationsinfrastrukturen im Wissenschaftssystem eine wichtige Rolle zukommt. Gleiches gilt insbesondere für den Kulturbereich, für den Bereich der Beteiligung von Laien an Forschungsprozessen (*Citizen Science*) sowie für die Datenproduzenten und Infrastrukturanbieter in Ämtern und Behörden. Forschende sollten in ihrer Doppelrolle als Produzenten und Nachnutzer von Daten ihre Interessen mit den Infrastrukturen abstimmen, und für Infrastrukturen als Serviceeinrichtungen gilt umgekehrt, dass sie die Abstimmung mit Forschenden zu suchen haben: Für beide empfiehlt der RfII die Schaffung eines Forums. Er empfiehlt darüber hinaus, auch bei der Realisierung der NFDI geeignete Beteiligungsformen zu etablieren, z. B. Beiräte oder *Communities of Practice*.

Breite Beteiligung

4.4.5 Den Belangen einer *Internationalisierung* (vgl. 4.8) soll die Vernetzung international aktiver Gremienvertreterinnen und -vertreter aus der deutschen Wissenschaft dienen. Forschungsförderer oder auch der RfII können hierfür ein geeignetes Forum etablieren.

Vernetzung
internationaler
Aktivitäten

4.5 NEUE BERUFSBILDER, STUDIENGÄNGE, AUSBILDUNGS- WEGE

Je stärker Prozesse an Hochschulen oder überhaupt forschenden Einrichtungen durch Digitalisierung bestimmt sind, desto dringlicher wird eine systematische Bündelung von Expertise. Datendienste, Datenmanagement oder die Ausgestaltung informationsinfrastruktureller Leistungen sind wissenschaftlich anspruchsvolle Aufgabenstellungen neuen Typs. Zugleich setzen Lehre und Forschung Professionalität und internationale Anschlussfähigkeit in solchen neuen, teils eher fachmethodischen, teils eher infrastrukturnahen oder dienst- und managementorientierten Kompetenzfeldern voraus. Die gegenwärtige Situation spiegelt diesen systematischen Zugang zu den neuen Herausforde-

rungen, der die Bereiche *Lehre, Forschung und Infrastrukturdienste/Services* gleichermaßen betrifft, noch nicht wider. Stattdessen überwiegt ein „Trend zum blumigen Berufstitel“⁷⁸ ohne Anbindung an dafür zuständige Ausbildungs- oder Studiengänge.

Der Rfll begrüßt die laufenden Selbstverständigungsprozesse der Hochschulen über geeignete neue Studiengänge und unterstreicht die Notwendigkeit eines nicht nur raschen, sondern auch differenzierten Wandels. Dieser muss sowohl der Vielfalt der Fächer und Forschungsmethoden gerecht werden als auch das breite Spektrum von Anforderungen berücksichtigen, das sich zwischen eher technischen und organisatorischen sowie fachwissenschaftlich-inhaltlichen Ausbildungsprofilen auftut. Die inzwischen zahlreichen im Wissenschaftssystem entstandenen Empfehlungen zur Entwicklung der Informationsinfrastruktur⁷⁹, die sich sowohl auf generelle Aufgaben als auch auf die Herausforderungen spezifischer Services beziehen, verlangen zu Recht von den Bildungsanbietern (Hochschulen, Fachverbände etc.) konkrete Programme und Umsetzungswillen.

Durchgehende Vermittlung von Kompetenzen in Fachstudiengängen

4.5.1 Der Ausbau von Informationsinfrastrukturen muss durch Stärkung der Informationskompetenz auf allen Ebenen flankiert werden. Informationsinfrastrukturen bestehen nicht nur aus IT-Strukturen und Softwarelösungen. Mindestens so wichtig sind die Erbringung von Services und fachnahe Methodenkompetenz. Der Begriff „Informationskompetenz“ muss weit gefasst werden, denn er führt unterschiedliche Teilfertigkeiten technischer, kommunikativer, sozialer bzw. organisationsbezogener Art und disziplinspezifischer Art zusammen.⁸⁰ Der Rfll empfiehlt, Module zur Vermittlung von Informationskompetenz und Datenmanagementkenntnissen in das gesamte Spektrum bestehender Studiengänge zu integrieren (also von den Ingenieurwissenschaften über die Medizin und die Rechtswissenschaft bis hin zu den Geistes- und Sozialwissenschaften), um kompetente künftige Forschergenerationen heranzubilden. Ebenfalls sind Fort- und Weiterbildungspfade – idealerweise international abgestimmt – für Wissenschaft und Verwaltung zu eröffnen. Die Ausgestaltung des Bildungsangebots im Spannungsfeld zwischen methoden- und fachwissenschaftlichen Disziplinen ist zu klären.⁸¹

Einrichtung spezialisierter Vollstudiengänge

4.5.2 Daneben sollten Vollstudiengänge eingeführt werden, die auf neue Berufsbilder wie z. B. den (digitalen) *Dokumentar*, den *Data Librarian*, den *Datenarchivar* oder den *Data Scientist* mit Spezialisierung in den jeweiligen Fächern

⁷⁸ Hanraths (2015) – Hacker und Missionare.

⁷⁹ Vgl. Übersicht zur Historie im Anhang B.

⁸⁰ Vgl. HRK (2012) – Hochschule im digitalen Zeitalter, S. 6.

⁸¹ Der Wissenschaftsrat hat in einem Positionspapier zwischen „Simulationswissenschaft“ (simulation sciences) und „simulationsgestützten Fachwissenschaften“ (simulation-based sciences) unterschieden. Vgl. WR (2014) – Positionspapier Simulation, S. 11. Eine ähnliche Unterscheidung könnte auch für die Datenwissenschaften sinnvoll sein.

hinführen. Zum Qualifikationsprofil dieser Berufsbilder hat die HRK formuliert, dass sie „sämtlich hohe Integrationsfähigkeiten [implizieren] und auf der Schnittstelle sowohl zwischen Wissenschaft und Infrastruktureinrichtungen als auch zwischen den Infrastruktureinrichtungen selbst [operieren], so dass die hergebrachte Trennung etwa zwischen Bibliothek und Rechenzentrum mit Blick auf jene Tätigkeit obsolet erscheint.“⁸² Aufgrund des absehbar hohen Bedarfs (vgl. 2.3) sollten neue Studiengänge an einer großen Zahl von Hochschulen von den bestehenden bibliotheks-, archiv- bzw. informationswissenschaftlichen Zentren in enger Kooperation mit der Informatik und den jeweiligen Fachdisziplinen vor Ort entwickelt und angeboten werden. Ebenso empfiehlt der Rfll, aktiv um Interessenten für die neuen Berufsbilder und Bildungspfade zu werben.

4.5.3 Aufgaben und Dienste des Forschungsdatenmanagements enthalten auch standardisierte und sich wiederholende Tätigkeiten, die nicht notwendigerweise durch akademisch ausgebildetes Personal erbracht werden müssen. Im Rahmen des dualen Ausbildungssystems wurden in den letzten Jahren forschungsdatennahe Ausbildungsgänge (z. B. Fachangestellte für Markt- und Sozialforschung) geschaffen. Derartige Ausbildungspfade müssen weiterentwickelt und gefördert werden. Durch den Rückgriff auf nicht-akademisch ausgebildetes Personal eröffnen sich Möglichkeiten einer weiteren und zugleich effizienten Professionalisierung forschungsdatennaher Aufgabenfelder in Infrastruktureinrichtungen.

Qualifiziertes
Personal über duale Be-
rufsbildung gewinnen

4.5.4 In der Ausgestaltung neuer Lehr- und Lernprozesse im Kontext der zunehmend digitalisierten Wissenschaft ist besonders darauf zu achten, dass diese chancenreichen neuen Berufsfelder und Karrierewege auch und in zunehmender Zahl von Frauen angenommen werden. Auch sollten Qualifizierungsoptionen für ältere Berufstätige sowie Zuwanderer angeboten werden, um die starke Nachfrage nach Fachkräften decken zu können. Die neuen akademischen und nicht-akademischen Berufsprofile werden sich sowohl innerhalb als auch außerhalb der Hochschulen einer großen Nachfrage erfreuen: Akademische Datenspezialistinnen und -spezialisten sind überall dort gefragt, wo es gilt, große oder komplexe Datenmengen zu managen und in komplexe Organisationszusammenhänge zu integrieren. Bei der Konturierung der Kompetenzfelder sollten daher einerseits anspruchsvolle datenbezogene Aufgaben auch aus nicht mit Forschung betrauten Feldern (im Gesundheitsbereich, in der öffentlichen Verwaltung, den Gedächtniseinrichtungen, im Kulturbetrieb etc.) Berücksichtigung finden. Andererseits ist mit einem ‚Wettbewerb um Köpfe‘ zu rechnen. Somit gilt es, qualifiziertes Personal in der Forschung zu halten. In der Konkurrenzsituation wird dies nur durch attraktive Ausgestaltung der neuen Berufswege gelingen können.

Wettbewerb um Köpfe

⁸² HRK (2012) – Hochschule im digitalen Zeitalter, S. 16.

4.6 NEUE „DATENKULTUR“

Mit der Digitalität von Forschungsdaten stellen sich vielfältige, in ihrer Art neue Herausforderungen an deren Qualitätssicherung. Neben der Qualität und Integrität von Datensätzen ‚an sich‘ spielen die Qualität des Datenmanagements und der Metadaten, die Einhaltung international gängiger Standards sowie die Qualität der Infrastruktur selbst eine wichtige Rolle.⁸³

Eine ausreichende Datenqualität über den gesamten Datenlebenszyklus zu sichern, erfordert eine gemeinschaftliche Anstrengung. Produzenten, Verarbeiter und diverse Nachnutzer von Forschungsdaten müssen hier jeweils Qualitätsverantwortung übernehmen. Dies bedingt eine neue Datenkultur der Offenheit und des Teilens (Data Sharing⁸⁴), aber auch der Verständigung auf Standards. Das Etablieren eines gemeinsamen Qualitäts-Diskurses aller Akteure ist ein Ziel, dessen Realisierung Zeit und zahlreiche Zwischenschritte erfordern wird. In einigen Bereichen des Wissenschaftssystems sind hierfür bereits kollaborative Strukturen im Aufbau, die Abstimmungsprozesse kanalisieren und mittelfristig die Durchsetzung gemeinsamer (Qualitäts-)Standards und Werkzeuge ermöglichen möchten.⁸⁵

Fachliche
Standards und Reviews
für Daten

4.6.1 Der RfII empfiehlt, dass Fachverbände, wissenschaftliche Vereinigungen, die DFG-Fachkollegien und ausgewiesene Infrastrukturbetreiber adäquaten, dabei aber auch hinreichend dynamischen fachlichen Standards und Qualitätskriterien große Aufmerksamkeit widmen. Fächerübergreifende Kooperationsmöglichkeiten sowie die Zusammenarbeit mit anderen Disziplinen und Forschungsformen müssen im Rahmen von Datenplattformen und virtuellen Forschungsumgebungen berücksichtigt werden.

Die Beschreibung der Datenqualität durch möglichst standardisierte und dabei auf Interoperabilität ausgerichtete Metadaten spielt hierbei eine entscheidende Rolle. Forschende sollten zunächst durch gute Dokumentationen und Berichte die Qualitätssicherung für ihre Forschungsdaten garantieren. Insbesondere müssen die in den Daten enthaltenen Unsicherheiten mitbeschrieben werden. Entsprechende Selbstverpflichtungen der Wissenschaft sind die Grundlage, um qualitätsgesicherte Forschungsdaten in den Repositorien anzubieten.

⁸³ Vgl. ausführliche Begriffsklärung für Datenqualität im Anhang A.

⁸⁴ Vgl. Begriffsklärung im Anhang A.

⁸⁵ Teils formieren sich Interessengruppen im Zusammenhang mit der Bildung verteilter Infrastrukturen, wie dies beispielsweise mit der Initiative ELIXIR für den Fächercluster der Lebenswissenschaften der Fall ist (vgl. 2.4, FN 47), teils setzen Infrastrukturanbieter Standards, wie im Falle der rund 30 akkreditierten Forschungsdatenzentren des RatSWD oder der internationalen Data Seal of Approval-Community (vgl. 2.5, FN 58).

In einem weiteren Schritt sollten geeignete *Review*-Verfahren konzipiert und in ausreichend vernetzten Communities durchgesetzt werden. Idealerweise werden Mischlösungen, die sowohl übergreifende als auch fachspezifische Qualitätskriterien berücksichtigen, aus den Communities heraus entwickelt und ausgestaltet. Dabei sollte frühzeitig der Kontakt zu (möglichen) Datennutzern und fachtypischen übergreifenden Verbänden gesucht werden (vgl. 4.2). Interoperabilität ist auch für Qualitätsstandards, die im ersten Schritt *bottom up* entstehen, ein wichtiges Ziel. Mittelfristig sollen sich für Forschungsdatenpublikationen verbindliche Mechanismen etablieren, die dem *Peer Review* zur Qualitätssicherung wissenschaftlicher Zeitschriften ähneln.

4.6.2 Der RfII empfiehlt, Datenmanagementpläne (DMP) als Planungsinstrumente verbindlich einzufordern und ihren breiten Einsatz im Forschungsprozess stärker voranzutreiben – nicht nur bei Drittmittelforschung, sondern als Teil guter wissenschaftlicher Praxis. Datenmanagementpläne sind bei Datenproduzenten aus Wissenschaft, Wirtschaft und Verwaltung als Qualitätsmanagementinstrument bereits vielfach im Einsatz. Bund und Länder sowie die Förderorganisationen können Standards für DMP in der Drittmittelforschung setzen. Hochschulen und Forschungseinrichtungen sollten darüber hinaus durch interne Policies das Datenmanagement institutionell verankern.⁸⁶ Gleichzeitig sollte die Umsetzung dieser Vorgabe und ihre Auswirkungen in verschiedenen Themenfeldern und auf unterschiedliche Forschungsformen durch Begleitforschung nachverfolgt werden, um ggf. einer wissenschaftsschädlichen Bürokratisierung entgegenwirken zu können.

Datenmanagementpläne als Planungsinstrument

4.6.3 Der RfII empfiehlt, der Nationalen Forschungsdateninfrastruktur (NFDI) Mechanismen einer regelmäßigen externen Qualitätssicherung (z. B. Monitoring und Evaluierung) beizugeben und damit für die wissenschaftlichen Nutzer und die Förderer Transparenz herzustellen. Der Prozessqualität der Repositorien und Dienste kommt besondere Bedeutung zu, da diese den gesamten Wertschöpfungsprozess begleiten müssen und eng in Entwicklung und Anwendung von Regelungen, Standards und Policies eingebunden sind. Für Produzenten und wissenschaftliche Nachnutzer wie auch für die an Daten interessierten Parteien aus Wirtschaft und Politik können Informationsinfrastrukturen als „Qualitäts-Dienstleister“ zur gefragten Anlaufstelle werden.

Externe Qualitätssicherung

Die Qualität von Repositorien und anderen Diensten bemisst sich auch an ihrer Fähigkeit, mit Daten entsprechend ihrer Relevanz zu verfahren. Daten, mögliche Nachnutzungsformen und die Relevanz von Daten lassen sich typisieren. Dimensionen für Relevanz sind beispielsweise wissenschaftlicher Nutzwert, wirtschaftliche Verwertungsperspektive, gesellschaftlich-politische Bedarfe an Daten oder langfristige kulturelle Bedeutung.

⁸⁶ Vgl. hier auch die handlungspraktischen Empfehlungen der HRK (2015) – Handlungsoptionen für Forschungsdatenmanagement und der LERU (2013) – Roadmap for Research Data.

Diese Dimensionen werden durch die Einschätzungen der jeweiligen Nutzergruppen maßgeblich geprägt (vgl. auch 4.3.3). Eine grobe Richtschnur für einen qualitätvollen Umgang mit Daten, z. B. zur nachvollziehbaren Abschichtung notwendiger Kuratierungsniveaus und -aufwände für Datensätze, wird für zahlreiche Informationsinfrastrukturen hilfreich sein und sollte bald zur Verfügung stehen.

Existierende Initiativen, die Qualitätsanforderungen für Repositorien sowie Kriterienkataloge zur Archivierung digitaler Objekte für die Wissenschaft entwickeln⁸⁷, sollten in die Entwicklung eines Rahmenwerks des Qualitätsmanagements eingebunden werden. Für die Qualität einzelner Dienste sind darüber hinaus Kriterien des *Quality of Service* (QoS) zu Aspekten wie Performanz, Verfügbarkeit, Konformität etc. zu definieren, die sich an aktuellen Standards vergleichbarer Informationsinfrastrukturen, z. B. im E-Government, orientieren. Bei allen Qualitätsmaßnahmen ist auf die Etablierung geeigneter Monitoring-Verfahren zu achten, die Evaluationen im System unterstützen.

4.7 NUTZUNGSANREIZE UND INNERWISSENSCHAFTLICHE AKZEPTANZ

Für Forschungsdatenmanagement-Initiativen ist das Verhalten der Forschenden ein erfolgskritischer Punkt. Für Datenproduzenten stehen oftmals die direkte Verwendung und das Teilen (Data Sharing) eigener oder auch fremder Daten im nahen Umfeld der eigenen Community im Vordergrund. Die Bedeutung des professionellen Managements, der Zugänglichkeit und der Langfristverfügbarkeit sind demgegenüber oft weniger präsent. Regeln guter wissenschaftlicher Praxis und Leitlinien zur Erstellung von Datenmanagementplänen wirken hier auf einen Kulturwandel hin. Gleichwohl bleibt eine Reihe von möglichen Hemmnissen für die Akzeptanz von Angeboten zum Umgang mit Forschungsdaten durch die einzelnen Forschenden. Jenseits der in 4.6 bereits angesprochenen Punkte bedarf es der Etablierung einer Forschungsdatenkultur, die auf Akzeptabilität und Attraktivität von digitalen Rahmenbedingungen gründet.

4.7.1 Die Hochschulen und Forschungseinrichtungen sollten ihren Forschenden Unterstützung und Service für das Forschungsdatenmanagement anbieten und auf nutzerfreundliche Strukturen hinwirken. Gerade in der Umbruchsphase kommt sowohl den Forschenden als auch den Forschungsfördernden Organisationen die Rolle von Beratern und „Brokern“ zu.

Unterstützung für
Forschungsdaten-
management vor Ort

⁸⁷ Vgl. die in 2.5 genannten Beispiele: Akkreditierungsverfahren des RatSWD, nestor-Siegel und Data Seal of Approval.

4.7.2 Nicht immer werden unter hohem Ressourceneinsatz etablierte Dienste in hinreichendem Maße genutzt. Häufiger Grund ist eine zu geringe Passfähigkeit zwischen angebotenen Werkzeugen und den oft sehr spezifischen oder einfach nicht hinreichend genau ermittelten Ansprüchen der Forschenden in puncto Nutzungsaufwand und Benutzerfreundlichkeit. Daher ist bei der Entwicklung und Evaluation von Diensten und Werkzeugen für das Forschungsdatenmanagement schon zu einem frühen Zeitpunkt auf Nutzerbedürfnisse zu achten. Deren Erfüllung durch die Infrastruktur sollte *ex post* geprüft werden.

Nutzerorientierte Gestaltung der Services

4.7.3 Mit digitalem Datenverkehr gehen typischerweise deutlich mehr rechtliche Unsicherheiten einher als beim herkömmlichen Datenaustausch der Fall ist. Regeln guter wissenschaftlicher Praxis und Datenmanagement-Policies erhöhen mittels Stichworten wie „Datenverantwortung“ oder auch „Datenqualität“ den Druck auf die einzelnen Forschenden – insbesondere auf die Leiterinnen und Leiter größerer Forschungsprojekte. Das Negativbild einer rechtlichen „Grauzone“ kann hier das Vertrauen in die der Wissenschaft grundgesetzlich zugebilligte und geschützte Freiheit des Umgangs mit dem für die Forschung essenziellen Datenmaterial überlagern. Der RfII empfiehlt dafür zu sorgen, dass in allen Organisationen möglichst vor Ort Information und Aufklärung als auch juristische Beratung verfügbar sind. Dieser in hohem Maße verhaltensrelevante Aspekt des digitalen Wandels muss auch auf der Ebene des gesamten Wissenschaftssystems angegangen werden. Geeignete Kommunikationsplattformen zum Austausch zwischen Forschungseinrichtungen, Forschungsförderern, wissenschaftsorientierten Stiftungen und der HRK sollten zeitnah geschaffen werden.

Rechtliche Beratung für Forschende

4.7.4 In Anlehnung an ähnliche Vorschläge empfiehlt der RfII die Herstellung einer Zitierbarkeit von Daten⁸⁸ und die Etablierung von Datennutzungs- und Datenzitierungsindizes, soweit Fächer über eine hierfür geeignete Publikationskultur verfügen. Diese Indizes sollen so ausgestaltet sein, dass sie Anreize setzen für die Bereitstellung sowie Nutzung von Forschungsdaten, und damit auch für deren Management. Einerseits kann die Verknüpfung eines Datensatzes oder einer Software mit den produzierenden Forschenden reputationsförderlich sein und positive Leistungsmotivation erzeugen. Andererseits kann die Nutzungsintensität von Datensätzen als ein Relevanzkriterium für die Entwicklung der sie anbietenden Infrastruktur gelten und forschenden Communities wie auch Förderern ggf. Entwicklungshinweise geben. Der RfII empfiehlt, aussagekräftige Indikatoren und mögliche Effekte ihrer Anwendung auf der Ebene der Fachgemeinschaften zu erörtern, z. B. über die DFG-Fachkollegien. Impulse für die Diskussion um die Weiterentwicklung von Anreizsystemen und die mitt-

Weiterentwicklung von Anreizsystemen

⁸⁸ Z. B. Data Citation Synthesis Group (2014) – Joint Declaration of Data Citation (Webseite) (FORCE 11) oder DataCite Metadata Working Group (2015) – Metadata Schema. Vgl. auch die Ausführungen zu Ermöglichungsstrukturen in 2.5.

lerweile kritisch reflektierte Praxis im Bereich der Bibliometrie⁸⁹ sind wünschenswert.

4.7.5 Der RfII empfiehlt Forschenden, die Organisation von Forschungsprozessen unter Bedingungen von Digitalität als Kompetenzfeld zu begreifen, dem mittel- und langfristig eine hohe Bedeutung zukommt. Jenseits von Aus- und Fortbildung (siehe 4.5) sollten sowohl Nachwuchswissenschaftlerinnen und -wissenschaftler als auch etablierte Forschende Informationskompetenz nicht als Zusatzqualifikation, sondern als zentralen Aspekt der Methodiken und der Methodendiskurse ihrer eigenen Fachlichkeit begreifen. Geeignete Maßnahmen, um diesbezüglich eine innerwissenschaftliche Verständigung voranzutreiben, könnten die Etablierung virtueller Diskussionsforen oder auch methodenorientierter Zeitschriften sein.

4.8 INTERNATIONALE FORSCHUNGS- UND INFORMATIONSDATEN- INFRASTRUKTURPOLITIK

Wissenschaft lebt vom Austausch von Daten auf internationaler Ebene und kann nur durch diesen Austausch international wettbewerbsfähig sein und zur Lösung globaler gesellschaftlicher Herausforderungen beitragen. Nationale Konzepte zum Umgang mit Forschungsdaten müssen daher das internationale Zusammenspiel sicherstellen und auch proaktiv mitgestalten. Deutschland ist als starker Mitgliedstaat der EU bereits in der Ausgestaltung eines europäischen Forschungsraums aktiv. Schon dieser Prozess wird durch Unterschiedlichkeit und Ungleichzeitigkeit der Entwicklungen in den einzelnen Mitgliedstaaten geprägt, auch was das Management von Forschungsdaten betrifft.⁹⁰ Für Abstimmungen auf globaler Ebene gilt dies in noch einmal gesteigertem Maße. Aber auch hier ist Deutschland in vielfältigen Foren und Organisationen aktiv.

Die Entwicklung eines Informationsinfrastruktursystems für Forschungsdaten in Deutschland unterliegt systemischen Rahmenbedingungen, die sich von denen anderer Länder unterscheiden. Hierzu gehören die verteilten wissenschaftspolitischen Verantwortlichkeiten und die Vielfalt der Finanzierungspfade. Top-down-Vorgaben, etwa eines für die öffentliche Finanzierung obligatorischen Datenmanagements, wurden bisher kaum realisiert. Dennoch sollte eine nationale Strategie auch durch Elemente verstärkt werden können, die sich in anderen wissenschaftlich führenden Nationen bewähren. Es gilt zu prüfen und ggf. durch Erprobung zu lernen, was unter deutschen Rahmenbedingungen umsetzbar ist. Dabei sollten traditionelle Stärken des deutschen Systems –

⁸⁹ Die Effekte von Zitierungsindizes sind zunehmend umstritten; vgl. unlängst kritisch Doove (2016) – Amsterdam Call for Action (Webseite).

⁹⁰ Vgl. Abschnitt 2.4 „Deutschland im internationalen Vergleich“ sowie ERAC (2016) – Open Research Data.

Dezentralität und leistungsfähige Vielfalt bei einem zugleich hohen Grad an Politikferne und wissenschaftlicher Selbstorganisation – nicht nur bewahrt und gepflegt werden, sondern durchaus in internationale Konzepte Eingang finden. Ebenso muss Forschungsdatenmanagement als Gestaltungsaufgabe gesehen werden, die aktiv zu verfolgen ist und für die deutsche Akteure in internationalen Strategiediskussionen Verantwortung übernehmen.

4.8.1 Um angesichts international sehr unterschiedlicher Gestaltungsansätze für den digitalen Umbruch in der Wissenschaft politisch bewertbare Entscheidungsvorschläge ableiten zu können, rät der Rfll dazu, Entwicklungstendenzen in anderen Ländern im interkulturellen Systemvergleich zu verfolgen, in Europa wie auch weltweit. Schlichte Vergleiche zwischen Lösungsansätzen, die staaten- bzw. länderspezifische Ausgangslagen außer Acht lassen, können freilich in die Irre führen. Der Rfll erachtet es daher als wichtig, vertiefte Kenntnisse über weltweit unterschiedliche nationale Forschungsdatenstrategien systematisch aufzubauen und mit den in Deutschland und im europäischen Kontext intendierten Entwicklungen zu vergleichen. Besonderer Aufmerksamkeit bedarf aus Sicht des Rfll das international unterschiedliche Management von sogenannten „disruptiven“ Entwicklungen, die als Innovationstreiber des stattfindenden Umbruchs angesehen werden. Um in einem so agierenden Umfeld mithalten zu können, müssen in Deutschland methodische Ansätze gefördert und in einen nationalen Fachdiskurs eingebracht werden. Die darauf basierenden Lösungen müssen international anschlussfähig und daher, eingedenk der internationalen Heterogenität, hinreichend flexibel sein. Angesichts der dynamischen Entwicklungen im Feld wird der Rfll im Rahmen seines Auftrags internationale Entwicklungen weiter beobachten, mit nationalen Zielsetzungen abgleichen und strategische Einschätzungen abgeben.

Lernen durch
Systemvergleich

4.8.2 Der Rfll sieht Bedarf an neuen Formen des Partizipationsmanagements, um deutsche wissenschafts- und technologiepolitische Teilhabe auf internationaler Ebene zu sichern. Die internationalen Entwicklungen werden zunehmend schneller und tiefergreifend durch neue und wechselnde Standards gestaltet. Diese werden *bottom up* und fachspezifisch durch die vernetzte Informationswelt formuliert, nicht selten aber auch *top down* durch Interessengruppen über formale Normen (z. B. ISO-Standards) durchgesetzt. Beide Ansätze stellen im Technologiebereich eine erfolgreiche Strategie dar, eigene Interessen durchzusetzen und davon zu profitieren. In den letzten zwei Jahrzehnten hat sich eine entsprechende Dynamik verstärkt. Der Rfll sieht es als erforderlich an, Vertreterinnen und Vertreter der deutschen Wissenschaft noch stärker zur Beteiligung an den entsprechenden Verfahren zu motivieren und sie dabei zu unterstützen. Nur eine intensivere aktive Teilnahme an solchen internationalen Gestaltungsprozessen wird es Forschenden aus Deutschland ermöglichen, die in Deutschland entwickelten Ansätze und Lösungsvorschläge bei den internationalen Abstimmungsverfahren einzubringen und zu vertreten. Ebenso sollten

Personen in Positionen mit Schlüsselfunktionen in internationalen Gremien im engen Austausch mit nationalen Gremien stehen und durch kommunikative Einbindung ihre potenziellen Nachfolger auf die Aufgaben in den internationalen Gremien vorbereiten. Die Gewinnung von Personen für derartige Aufgaben muss ein gemeinsames Anliegen aller mit Forschungsorganisation befassten Akteure sein (vgl. auch 4.4.5). Die Wissenschaftsorganisationen könnten dazu beitragen, dass die Tätigkeit in internationalen Gremien auch von der Community besser wertgeschätzt wird als bislang.

Praxisnah international weiterbilden

4.8.3 In Deutschland wird viel über die Nachteile und Gefahren des digitalen Umbruchs in Wissenschaft, Forschung und Entwicklung diskutiert. Dass gleichwohl die dringende Notwendigkeit besteht, neue technologische Möglichkeiten auch in ihren internationalen Kontexten zu verstehen, um den digitalen Wandel mit all seinen Chancen und Risiken proaktiv zu gestalten, ist demgegenüber weniger präsent – in Forschung, Hochschullehre und Bildungsbereich selbst auf Leitungsebenen. Neben international abgestimmten Aus- und Weiterbildungswegen auf der nationalen Ebene (vgl. 4.5) regt der RfII daher an, im Sinne einer Befähigung zum Wandel auch im Rahmen internationaler Initiativen und Informationsinfrastrukturen praxisorientierte Weiterbildungsangebote auszuarbeiten und die praxisnahe Forschung an und mit existierenden Dateninfrastrukturen durch Stipendien und Projektmittel zu fördern.

Internationale Kooperation auf Augenhöhe

4.8.4 Internationalität erfordert organisatorische Schnittstellen, nicht zuletzt aufgrund deutscher Binnenkomplexität. Voraussetzung für eine erfolgreiche Internationalisierung ist, dass nationale Informationsinfrastrukturen nach außen als starke Kooperationspartner agieren können. Der RfII sieht insbesondere eine Verpflichtung, gemeinsame Strukturen im europäischen Forschungsraum zu unterstützen und Barrieren abzubauen. Aktive internationale Beteiligungen werden erleichtert durch klar formulierte, konsenterte und implementierte Vorstellungen über Governance, technische und organisatorische Anschlussfähigkeit, nachhaltige Finanzierungsmöglichkeiten, Qualitätskriterien und Nutzerorientierung. Aus Sicht des RfII bedarf es weiterer, konsequent gesetzter Anreize zur Konsolidierung nationaler Selbstorganisation im Sinne einer Interessenvertretung und „Stimme“ für Informationsinfrastrukturen – mit dem Ziel einer handlungsfähigen Governance, um mit ihren internationalen Pendanten auf Augenhöhe kooperieren zu können. Vorbilder sind hier die sehr aktiven Akteure in den europäischen ESFRI-Projekten und beim Aufbau leistungsfähiger Rechenressourcen (z. B. Gauß-Zentrum für Supercomputing, Gauß-Allianz). Forschungspolitik sollte analoge Entwicklungen sowohl durch Fördermaßnahmen als auch durch Begutachtungen unterstützen.

4.9 NUTZUNG UND VERWERTUNG

Daten werden vielfach als Rohstoff des Informationszeitalters bezeichnet. Bei Forschungsdaten ist der rechtliche Status häufig unklar, vor aneignender bis hin zu unrechtmäßiger Nutzung sind sie nicht grundsätzlich geschützt. Dies gilt insbesondere, wenn Forschende Daten selbst weitergegeben haben, etwa an kommerziell interessierte Online-Netzwerke oder in eine kommerzielle Cloud (vgl. auch 2.2). „Offenheit“ von Forschungsdaten schafft daher nicht nur Wege des Teilens und der wissenschaftlichen Nutzung (Open Data, Open Science⁹¹), sondern auch Möglichkeiten einer wirtschaftlichen Verwertung dieser Daten.

Der kommerzielle Zugriff auf Forschungsdaten kann unbedenklich und gewollt sein und unter Berücksichtigung der wissenschaftlichen Belange (sowie Datenschutz, Urheberrecht, Datensicherheit) transparent geregelt werden. Datenzugriff kann aber auch im Netz unter rechtlich unklaren Bedingungen oder unbemerkt erfolgen. Im Extremfall kann eine nachgelagerte Privatisierung (also die Aneignung „offener“ Daten durch wirtschaftliche Akteure im internationalen Rechtsraum) eine weitere offene wissenschaftliche Nutzung oder eine Verwertung auf klaren Rechtsgrundlagen gerade verhindern (z. B. Patente).

Auch im Interesse der deutschen Wirtschaft sind Wissenschaft und Politik aufgerufen, auf die Herausforderung verlässlicher, transparenter, den Wert von Wissen schützender und der Erzeugung neuen Wissens förderlicher Regulierung die richtigen Antworten zu finden. Denn Datenwirtschaft ist globalisiert, Wertschöpfung wird vielfach in Ländern betrieben, die andere Nutzungsregeln haben.

4.9.1 Der innerwissenschaftliche Umgang mit Forschungsdaten an der Schnittstelle Wissenschaft/Wirtschaft bedarf aktiver Gestaltung. Aus Sicht des Rfll sind die Wissenschaftsorganisationen, aber auch die forschenden Einrichtungen gefordert, verbindliche Verhaltensregeln zu etablieren, um den unregulierten Datenfluss im Rahmen der Inanspruchnahme von Diensten privater Anbieter von Software, Kommunikations- und Speichermöglichkeiten zu gestalten. Alle Beteiligten sollten entschlossen auf einen Kulturwandel hinwirken, der das Thema der Offenheit für wissenschaftliche Nutzung mit dem Gedanken einer Datenverantwortung verbindet. Policies, Selbstverpflichtungen sowie eine sich gegen Aneignung aussprechende Lizenzierung (etwa, allerdings nicht rechtsverbindlich, als Creative Commons) sind mögliche Werkzeuge. Gebraucht wird ebenso ein Diskurs, der alle an Forschung beteiligten Akteure – bis hin zu Studierenden – für das Doppelgesicht des Schlagwortes Offenheit sensibilisiert.

Daten-
verantwortung

⁹¹ Vgl. Begriffsklärung im Anhang A.

4.9.2 Die Möglichkeit der Publikation von Forschungsdaten, also deren digitaler Bereitstellung (gesondert oder im Rahmen von Aufsätzen), stellt eine der großen, begrüßenswerten Chancen des Digitalzeitalters dar. Gebraucht werden jedoch geeignete Mindeststandards der Ausgestaltung von Datenpublikationen, welche die Zugänglichkeit und Nachnutzung der publizierten Daten kurz- und mittelfristig sichern (vgl. 4.6.1). Ebenso wichtig ist die Schaffung juristischer Standards für den wachsenden Bereich der sozialen Netzwerke. Der RfII befürwortet hier eine rechtliche Klärung sowie den Aufbau wissenschaftseigener Angebote, die auch deutschem Recht unterliegen (so etwa im Cloud-Bereich, vgl. 2.5 bzw. 4.2.6).

4.9.3 Werden Forschungsdaten publiziert, bedarf es forschungsfreundlich ausgestalteter Lizenzmodelle, die die Nachnutzung und im Falle von Texten auch die Anwendung von Textmining-Verfahren nicht verhindern. Zur rechtspolitischen Ausgestaltung eines zukunftsfähigen deutschen Urheberrechts verweist der RfII auf die Erklärung der Allianz der Wissenschaftsorganisationen vom 1. Dezember 2014.⁹²

4.10 MEHR WISSEN ÜBER FORSCHUNGSDATEN

Um Fehlsteuerungen auf dem Weg zu einem balancierten Ökosystem für Forschungsdaten zu vermeiden, wird gesichertes Wissen über die sich wandelnden Realitäten im Wissenschaftssystem benötigt. Aktuell orientieren sich Szenarien für das Forschungsdatenmanagement an Beispielen wissenschaftlicher Vorreiter-Communities, die sehr aktiv datenbasiert forschen. Darüber hinaus geben vereinzelt Studien Aufschluss über Einstellungen von Forschenden zu Paradigmen wie Open Science oder Data Sharing⁹³ und ihre Wünsche an eine unterstützende IT-Umgebung.⁹⁴ Insgesamt fehlt es jedoch an Daten, Fakten und Modellen zur Beschreibung der in Deutschland sehr fragmentierten und veränderlichen Landschaft im Bereich des Forschungsdatenmanagements. Schon zu Art und Anzahl, erst recht aber zu Nutzung/Erfolg von Infrastruktur-Initiativen, Betriebskosten und Investitionen sowie weiteren elementaren Kennzahlen sind kaum Angaben verfügbar.⁹⁵ International ist die Orientierung

⁹² Vgl. Allianz der Wissenschaftsorganisationen (2014) – Zum Urheberrecht.

⁹³ Vgl. auch Begriffsklärungen im Anhang A.

⁹⁴ Z. B. KE- Knowledge Exchange (2014) – Sowing the seed, die österreichische Studie von Bauer et al. (2015) – Forschende und ihre Daten oder der Beitrag von Fecher et al. (2015) – Reputation Economy.

⁹⁵ Zur fehlenden institutionellen „Landkarte“ vgl. Meyer-Doeringhaus/Neuroth (2015) – Stärkung Informationskompetenz, S. 81. Das Verzeichnis re3data.org listet immerhin knapp 250 Datenrepositorien mit deutscher Beteiligung, die DFG-Datenbank RIsources rund 200 Informationsinfrastrukturen auf (Stand April 2016).

noch schwieriger. Typologien für Vergleiche, Planung und Prognostik fehlen nach Kenntnis des RfII jenseits grober Schematisierungsversuche⁹⁶ ganz.

4.10.1 Der RfII empfiehlt, eine nationale Datenbasis zu den Belangen von Forschungs- und Informationsinfrastrukturpolitik zu schaffen. Verknüpft mit der derzeit in Weiterentwicklung befindlichen Forschungsberichterstattung über den Kerndatensatz Forschung⁹⁷ sollten dazu auch weitere Digitalisierungsindikatoren für Wissenschaft und Gesellschaft Berücksichtigung finden. Entstehende Systeme zur Erfassung von Daten über das Infrastruktursystem sollten international anschlussfähig sein. Über die nationale Ebene hinaus hält der RfII ein systematisches, vorzugsweise jährliches Monitoring von Entwicklungen in den europäischen Nachbarstaaten und auf der Ebene globalisierter Forschungs- und Technologiepolitik sowie Datenwirtschaft für erforderlich (vgl. 4.8.1).

4.10.2 Relevant sind insbesondere ein Verständnis des Verhaltens forschender Akteure in der ganzen Breite der Methoden und Forschungsformen, Studien zur rechtspolitischen und ökonomischen Bewertung von Organisationsmodellen sowie fundierte Risikobewertungen im Bereich von Datenschutz und Datensicherheit. Hierzu regt der RfII dringend an, gezielt gesellschaftswissenschaftliche Begleitforschung zu initiieren.

Datenbasis für
Forschungs- und
Infrastrukturpolitik

Begleitforschung zu
Digitalisierung in der
Wissenschaft

4.11 DATENSCHUTZ

Datenschutz ist ein rechtspolitisches Schlüsselthema des Digitalzeitalters. Forschungsdaten sind hier in gegenläufiger Weise von Interesse. Zum einen benötigt Wissenschaft auch Zugriff auf geschützte Daten. Ohne die Arbeit auch mit personenbezogenem, teils individualisiertem Wissen sind Forschung und wissenschaftlicher Fortschritt in Bereichen von Gesundheit, Bildung, Arbeit, Umweltschutz und öffentlicher Daseinsvorsorge nicht möglich. Zum anderen hat Wissenschaft aber auch Datenschutz zu gewährleisten. Im deutschen Recht geschieht dies auf der Basis von informationeller Selbstbestimmung und Persönlichkeitsrechten aus Art. 2 i.V.m. Art. 1 GG, international wird Datenschutz im Sinne von *Privacy* rechtlich etwas anders gefasst.

Deutsche Datenschutzbestimmungen sind im internationalen Vergleich streng und tragen der Spannung zwischen Schutzbedürfnis und gerechtfertigtem Forschungsinteresse nach Art. 5 GG Rechnung. Dennoch erwachsen durch niedrigschwellig auf Geschütztes zugreifende Digitaltechnik, durch ein enorm steigen-

⁹⁶ Vgl. etwa die „Data Pyramid“ der EU-High Level Expert Group on Scientific Data in: European Commission (2010) – Riding the wave, S. 18.

⁹⁷ Zur Erfassung von Forschungsinfrastrukturen in der Forschungsberichterstattung vgl. auch WR (2016) – Kerndatensatz Forschung, u. a. S. 43.

des Aufkommen an personenbezogenen oder auf Personen zurückführbaren Daten, durch die in manchen Feldern entstehende Unmöglichkeit von Anonymisierung⁹⁸ und durch die Globalität des Datenaustauschs neue Herausforderungen für den Datenschutz. Ebenso müssen Problemstellungen der *Datensicherheit*, also technische wie auch forensisch-kriminologische Fragen, mit dem Thema *Datenschutz* stärker verbunden werden (vgl. hierzu 4.12).

Der RfII weist in diesem Positionspapier vorerst nur auf die Bedeutung des Themas hin. Auf europäischer Ebene ist mit der EU-Datenschutzgrundverordnung⁹⁹ ein neuer rechtlicher Rahmen für die Mitgliedstaaten entstanden, der mit einer Übergangsfrist von zwei Jahren auch eine Anpassung deutscher Datenschutzbestimmungen erfordert. Die Verordnung enthält neben allgemeinen Vorschriften auch Bestimmungen zur Datenverarbeitung zu wissenschaftlichen Zwecken. Die Umsetzung auf nationaler Ebene muss genutzt werden, um die wachsenden Möglichkeiten und Bedarfe wie auch die Chancen und Risiken in den Blick zu nehmen, die sich aus den technologischen und methodischen Entwicklungen im Bereich der Datenanalyse künftig ergeben werden. Der RfII legt hierzu demnächst Empfehlungen in Form eines „Datenschutz-Kompass für die Forschung“ (Arbeitstitel) vor.

4.12 DATENSICHERHEIT

Digitale Forschungsdaten sind verletzlich. Datensicherheit ist daher Grundlage nicht nur der geordneten wissenschaftlichen Arbeit, sondern auch elementar für die Qualität sowie den wissenschaftlichen, ökonomischen und gesellschaftlichen Wert von Forschungsdaten. Zu unterscheiden sind *Safety* (Betriebssicherheit) und *Security* (Sicherheit vor Angriffen). Sind vernetzte Systeme der Datenhaltung technisch nicht robust ausgelegt oder können sie angegriffen werden (Stichwort Schadsoftware, „Hacking“ etc.), läuft Wissenschaft in womöglich neue Gefährdungslagen hinein.

Forschungsdatensicherheit besitzt aus Infrastrukturperspektive zwei maßgebliche Aspekte: die prinzipiell *sichere Speicherung* (auch langfristig) von Forschungsdaten in hierfür hinreichend leistungsfähigen, aber auch mit hinreichender Redundanz ausgestatteten Infrastrukturen einerseits sowie andererseits spezielle technisch-organisatorische Maßnahmen der *Cybersicherheit*, die missbräuchliche Zugriffe, Spionage und Sabotage verhindern.

⁹⁸ So sind etwa im biomedizinischen Bereich auch die direkten Verwandten einer Person, deren Gewebeprobe analysiert wird, biologisch identifizierbar.

⁹⁹ Council of the European Union (2016) – Standpunkt zur Datenschutz-Grundverordnung.

4.12.1 Der RfII empfiehlt den verantwortlichen Akteuren, technisch-organisatorische Maßnahmen zur Datensicherheit bei der Entwicklung von Informationsinfrastrukturen deutlich stärker als dies bisher der Fall ist, in den Fokus zu nehmen. Angesprochen sind hier neben der NFDI nicht nur Rechenzentren und Informationsinfrastrukturen generell, sondern auch Kostenträger und Forschungsförderer. Bei der Bewertung von Infrastrukturprojekten, insbesondere bei der Evaluation bzw. Akkreditierung zur Verstetigung von Diensten und Lösungen (vgl. 4.1) sind Datensicherheit und Datenschutz gleichrangig zu behandeln.

Datensicherheit im
Betrieb priorisieren

4.12.2 Wie der Datenschutz ist auch die Datensicherheit in hohem Maße abhängig vom Verhalten der Beteiligten in Forschung und Lehre. In dieser Hinsicht sind alle wissenschaftlichen Einrichtungen aufgerufen, beim Diskurs über Datenschutz auch Aspekte der Datensicherheit zu berücksichtigen.

Datensicherheit bei
den Forschenden ins
Bewusstsein rücken

4.13 FINANZIERUNG UND NACHHALTIGKEIT

Rahmenbedingungen für ein zukunftsfähiges Management digitaler Forschungsdaten in Deutschland zu schaffen, ist mit personellen, kommunikativen und finanziellen Aufwänden verbunden. Der RfII kann den Gesamtaufwand der empfohlenen Neugestaltung seriös nicht abschätzen bzw. beziffern. Die Größenordnung dürfte aber erheblich sein – zumal, zieht man neben der Komplexität der Aufgaben und der Zahl der Akteure auch die unvermeidliche Zeitdauer der erforderlichen Reorganisations- und Qualitätssicherungsprozesse in Betracht.

Die Anstrengungen richten sich allerdings darauf, eine kritische Schwelle der Entwicklung des Gesamtsystems Wissenschaft zu überwinden. Es geht also um eine Übergangsphase, in der gemeinsame Aktivitäten für einen außergewöhnlichen und in vieler Hinsicht auch einmaligen (sowie im Ergebnis weichenstellenden) Aufbau- und Transitionsprozess gefordert sind. Zugunsten neuer Produktivität von Forschung und Betrieb sowie internationaler Wettbewerbsfähigkeit müssen Wissenschaft und Politik wichtige Kernprozesse der inzwischen durch Digitalität geprägten Forschung neu gestalten. Es gilt, Potenziale zu heben, Effizienz zu steigern und drohende Fehlentwicklungen (Stichworte: Kleinteiligkeit, Projekt- und Insellösungen, Personalprobleme, fehlende oder heterogene Standards, Verlust von Nutzungsrechten, Datenverluste etc.) zu vermeiden. Hierbei entstehen Investiv- und Transitionskosten, die den Normalbetrieb von Forschung, Lehre und Infrastrukturentwicklung übersteigen.

Ziel der durch den RfII ausgesprochenen Empfehlungen und namentlich der Empfehlung zur Etablierung einer NFDI ist es, dem deutschen Wissenschaftssystem eine weitere quantitative und qualitative Entwicklung auf höchstem Niveau zu eröffnen. Ohne zusätzliche Ressourcen wird die zur Erreichung dieses Zieles dringend gebotene Umgestaltung nicht realisierbar sein. Ebenso sind

hinsichtlich einer wissenschaftsgemäßen Nachhaltigkeit des Datenmanagements hinreichend niedrigrschwellige, aber die Nutzung knapper Ressourcen sinnvoll steuernde Formen einer Umlage von Kosten für datenbezogene Services zu entwickeln.

Finanzierungsplanung mittels Roadmap

4.13.1. Der RfII hält Sonderanstrengungen von Wissenschaft und Zuwendungsgebern für erforderlich, um einen klug gesteuerten und auch durchgreifend wirksamen Transitionsprozess für das deutsche Forschungsdatenmanagement realistisch zu finanzieren. Er betont hierbei die Bedeutung der gemeinschaftlichen Dimension. In der Schwellensituation wird daher eine Finanzierungsplanung („Nationale Roadmap“) empfohlen, die über eine mehrjährige Übergangsphase kraftvolle Sondermaßnahmen aller Beteiligten und starke Anreize vorsieht. Aus Sicht des RfII ‚wuchert‘ sonst der Status quo ungesteuert in eine fehlentwickelte, die Leistungsfähigkeit der Wissenschaft in Deutschland schwächende und auch durch Ineffizienz geprägte kostenintensive Lage hinein.

Aufbau- und Transitionskosten, welche die normalen Investitionsspielräume im Wissenschaftssystem übersteigen, fallen aus Sicht des RfII insbesondere in folgenden fünf Feldern an:

- Aufbau der NFDI, Policy-Entwicklung *top down* (Projekt/ggf. Institutionalisierung von Kernaufgaben);
- Kommunikations- und Koordinationsprozesse auf der Ebene von Fachgemeinschaften und Disziplinen (Anreiz/ggf. Projekte);
- Reorganisationsprozesse bezüglich vorhandener Dienste (Anreiz/ggf. Verstetigungslösungen);
- Aus- und Fortbildungsaufgaben bezüglich der Handhabung von Daten im Forschungskontext (Anreiz/ggf. Projekte);
- Kommunikation und Monitoring als Querschnittsaufgabe (Projekte/ggf. Institutionalisierung von Kernaufgaben).

Beitragsmodelle für gemeinschaftliche Finanzierung

4.13.2. Für ein nachhaltiges deutsches Forschungsdatenmanagement sollten maßgeschneiderte Beitragsmodelle gefunden werden. Der RfII empfiehlt hier eine Experimentierphase, in welcher Anreize gesetzt werden, Formen der Kostenumlage zu entwickeln und zu erproben (wobei ggf. auch in sinnvoller Weise, z.B. nach Communities, Nutzergruppen o.ä. differenziert werden kann).

Grundsätzlich sind Beiträge sinnvoll, um den Wert von Diensten deutlich zu machen und Anreize für ressourcenschonendes Verhalten zu setzen. Allerdings dürfen durch Kostenumlagen keine demotivierenden Effekte entstehen. Auch darf der nachhaltige Umgang mit Forschungsdaten keiner kommerzialisierten Betrachtung unterliegen, die kostendeckende Preise oder gar Marktpreise zum Maß hat. Dies würde den wissenschaftlichen Wettbewerb und überhaupt die spezifischen Mehrwerte des gesellschaftlichen Systems Wissenschaft und der von ihm erbrachten Forschungsleistungen zerstören.

4.13.3. Monitoringmaßnahmen und Formen der Moderation sollten den Transitionsprozess (unter Wirksamkeits- und Effizienzgesichtspunkten) begleiten. Sie sollten die Frage der Auswirkung von Finanzierungsmodellen für Forschungsdatendienste auf das Gesamtsystem prüfen sowie die mittelfristige, *lernende* Ausgestaltung eines dynamischen Zustandes für eine *digitale* Wissenschaft der Zukunft begleiten. Lernprozesse können die Umschichtung von Ressourcen oder auch die Verschiebung von Aufgaben zwischen Einrichtungen betreffen.

Im Ergebnis müssen im Rahmen wissenschaftseigener Geschäftsmodelle Basisdienstleistungen auch für datenintensive Forschung in jedem Fall sichergestellt sein. Neben der Gewinnung von Daten ist die intensive Nutzung und insbesondere die forschende Nachnutzung vorhandener Daten ein zentrales Ziel auch für die Ausgestaltung von nachhaltigen Nutzungskonditionen. Ebenso ist für Motivation Sorge zu tragen, also dafür, dass möglichst viele Forschende sich für qualitativvolles Forschungsdatenmanagement einsetzen und ihre Daten in ein übergreifendes, qualitätsgesichertes, deutschen Rechtsregeln unterliegendes *Forschungsdaten-Ökosystem* einbringen.

5 AUSBLICK UND PRIORISIERUNG DER EMPFEHLUNGEN

Der Rat für Informationsinfrastrukturen (RfII) spricht nach einer Bestandsaufnahme umfassende Empfehlungen aus, die aufzeigen sollen, wie in Deutschland die Weichen für eine wesentlich auf digital verfügbaren Forschungsdaten ausgerichtete Wissenschaft zu stellen sind. Zahlreiche Aspekte des Forschungsdatenmanagements werden in der Folge noch genauerer Ausgestaltung bedürfen.

Der RfII verbindet seine Vorschläge mit der Botschaft an Politik und Gesellschaft, dass Digitalisierung zwar Veränderungen mit sich bringt, die Forschung im Land und weltweit aber befördert und zum Vorteil aller Bürger vorwärts entwickeln kann. Zu betonen ist noch einmal, dass es in komplexer Lage des Zusammenwirkens bedarf: Akteure in der Wissenschaftspolitik und Akteure auf Seiten der Selbstorganisation der Wissenschaft tragen hier in ihren Bereichen jeweils Verantwortung dafür, dass es zu gemeinsamem Handeln kommt. Entscheidend muss es sein, den Wandel gut zu gestalten, Deutschland in einer Situation des globalen wissenschaftlichen Wettbewerbs strategisch zu positionieren, öffentliche Mittel möglichst effizient einzusetzen und ein maximales Wertschöpfungspotenzial für Wissenschaft, Wirtschaft und Gesellschaft nachhaltig zu realisieren.

Die Erwartungen, mit welchen der RfII an Entscheider in Wissenschaft und Politik herantritt, lassen sich in fünf Punkten zusammenfassen. Sie benennen dringliche Handlungsfelder, die nicht nacheinander, sondern sofort und gleichzeitig angegangen werden können und sollten. Der RfII erachtet sie als gleichermaßen wichtig:

- Fördermechanismen anpassen
- Effizienz und Koordination durch eine (verteilte) nationale Infrastruktur sicherstellen
- „Forschungsdatenkultur“ befördern
- Übergreifendes Monitoring und Qualitätssicherung einrichten
- Personalentwicklung auf allen Ebenen betreiben.

Die Empfehlungen des RfII richten sich an die Gemeinsame Wissenschaftskonferenz des Bundes und der Länder, jedoch auch an andere Akteure, namentlich auch solche aus der Wissenschaft selbst, die den anstehenden Wandel tragen und gestalten müssen. Der RfII hat sich daher entschieden, die abschließende Priorisierung der von ihm empfohlenen Punkte nach Adressatengruppen zu ordnen.¹⁰⁰

¹⁰⁰ Für vertiefende Informationen sei außerdem auf die vier, das Positionspapier ergänzenden Anhänge hingewiesen (zu Begriffsklärungen, zur Geschichte der Informationsinfrastrukturpolitik in der Bundesrepublik Deutschland, zu beispielhaften „Szenarien“ der Forschungsdatenentstehung sowie zu Daten und Fakten).

Tabelle 1 fasst Empfehlungen hoher Priorität zusammen, die sich gleichermaßen an Bund und Länder wie die Wissenschaftsorganisationen richten. Sie zeigen die „großen Linien“ auf, die gemeinschaftlich zu gestalten sind. In *Tabelle 2* sind weitere Empfehlungen zusammengefasst, die sich primär an die Wissenschaftsorganisationen mit ihren Repräsentanten und Mitgliedsorganisationen richten. Sie betreffen Aufgaben, die im Rahmen der wissenschaftlichen Selbstorganisation zu lösen sind. *Tabelle 3* gibt eine Übersicht von Empfehlungen an Forschende, ihre Fachgesellschaften und Leitungsebenen. Diese betreffen Aufgaben, die durch diese Adressatengruppe umgesetzt, getragen und gelebt werden müssen oder die einer primär innerwissenschaftlichen Lösung bedürfen.

Tabelle 1: Empfehlungen für Bund, Länder und Wissenschaftsorganisationen

Nr. im Text	Empfehlungen mit höchster Priorität
4.1.1, 4.1.2	Phasenmodell für die Entwicklung von Informationsinfrastrukturen – Planbarkeit und Mindeststandards sicherstellen – geordnete Übergänge in geeignete Trägerschaft über unabhängige Begutachtungen organisieren
4.2.1	Aufbau einer Nationalen Forschungsdateninfrastruktur (NFDI) – Kompetenzen bündeln und Grundversorgung mit Services für das Forschungsdatenmanagement schaffen
4.2.3	Arbeitsteilige Organisation von Services in Verbundstrukturen – übergreifende Infrastruktur- und Kompetenzzentren etablieren
4.2.7, 4.4.1, 4.13.1	Gemeinsame Sonderanstrengungen von Bund, Ländern und Wissenschaftsorganisationen – Roadmap als verbindliches Koordinationsinstrument und zur Finanzierungsplanung für die NFDI
4.3.4, 4.3.5, 4.4.1	Nachhaltigkeit und Langzeitverfügbarkeit – Verstetigungsoptionen und Verantwortlichkeiten im Bereich der Langzeitarchivierung klären
4.6.2	Datenmanagementpläne als Planungsinstrumente im Forschungsprozess einsetzen

Nr. im Text	Empfehlungen mit hoher Priorität
4.2.2	Technische Ausstattung, Support und Methodenentwicklung im Zusammenhang sehen und planen
4.5.3	Forschungsdatennahe Aufgabenfelder auch durch den Rückgriff auf Ausbildungsberufe professionalisieren – forschungsdatennahe Ausbildungsgänge im dualen Ausbildungssystem fördern
4.6.3	Der NFDI ein externes Qualitätssicherungssystem begeben
4.9.3	Forschungsfreundlich ausgestaltete Lizenzmodelle für Forschungsdaten (inkl. Texten) entwickeln
4.8.1, 4.10.1, 4.10.2, 4.13.3	Steuerungswissen durch Monitoring und Begleitforschung erhöhen
4.13.2	Formen der Kostenumlage entwickeln und erproben, um maßgeschneiderte Beitragsmodelle für die NFDI zu finden

Tabelle 2: Weitere Empfehlungen für Wissenschaftsorganisationen

Nr. im Text	Empfehlungen mit höchster Priorität
4.1.3	Konzept für regelmäßige Begutachtungen von Forschungsdateninfrastrukturen entwickeln und etablieren
4.3.2, 4.3.3	Leitlinien, Datenmanagementpläne und Datenkuratierung in der Langzeitarchivierung standardisieren
4.4.3 4.4.4	Kriterien und Zertifizierungen für die Vernetzung von Informationsinfrastrukturen; Strukturen und Bedarfe im Hinblick auf eine nationale Forschungsdateninfrastruktur – Verständigung auf nationaler Ebene erzielen
4.5.1, 4.5.2, 4.8.3	Aus- und Weiterbildungsangebote sowie spezialisierte Vollstudiengänge einrichten – Informations- und Datenmanagementkompetenz in der Breite und auf allen Ebenen vermitteln – praxisorientierte Weiterbildung im Rahmen internationaler Initiativen und Informationsinfrastrukturen anbieten
4.6.1, 4.6.2	Standards und Qualitätskriterien durch Fachvertretungen und Infrastrukturbetreiber befördern und verbreiten – Datenmanagementpläne als Instrument guter wissenschaftlicher Praxis einsetzen
4.7.1 4.7.2	Unterstützung und Service für das Forschungsdatenmanagement anbieten und auf die Erfüllung der Nutzerbedürfnisse hinwirken
4.7.5	Organisation von Forschungsprozessen unter Bedingungen von Digitalität im Personalbereich als Kompetenzfeld begreifen
4.8.4	Gemeinsame Strukturen im europäischen Forschungsraum schaffen und Kooperation auf Augenhöhe erreichen
4.9.1	Verbindliche Verhaltensregeln etablieren, um den Datenfluss zwischen Wissenschaft und Wirtschaft zu gestalten – Ziel: Kultur, die Offenheit und Datenverantwortung miteinander verknüpft

Nr. im Text	Empfehlungen mit hoher Priorität
4.2.4	Übergreifende Dienste, vereinheitlichende Datenbeschreibungen und universelle Zugänglichkeit für die NFDI entwickeln
4.2.5 4.2.6	Funktionale Differenzierung von Einrichtungen innerhalb der NFDI – als Infrastrukturanbieter durch Methodenkompetenz und überzeugende Services Anreize für die Nutzung setzen
4.4.5, 4.8.2	International aktive Gremienvertreterinnen und -vertreter aus der deutschen Wissenschaft untereinander vernetzen – deutsche wissenschafts- und technologiepolitische Teilhabe sichern
4.5.4	Für das Forschungsdatenmanagement qualifiziertes Personal durch attraktive Ausgestaltung der neuen Berufswege in der Forschung halten
4.7.4	Datennutzungs- und Datenzitierungsindizes etablieren, soweit fachlich angemessen
4.9.2	Wissenschaftseigene Angebote für Datenpublikationen und im Bereich der sozialen Medien aufbauen, die auch deutschem Recht unterliegen

Tabelle 3: Empfehlungen für Forschende, Fachgesellschaften und Leitungsebenen

Nr. im Text	Empfehlungen
4.2.3, 4.2.5, 4.4.2	Möglichst in Verbundstrukturen arbeiten, die in die NFDI eingebettet werden können – existierende Services nutzen – Anbindung an größere Infrastrukturzentren suchen
4.3.1	Fachdiskurs über die Differenzierung zwischen projektlaufzeitnaher Speicherung sowie über deutlich länger zugeschnittene Archivierungsfristen führen
4.3.2, 4.3.3	Umgang mit Daten und die entlang des Datenlebenszyklus benötigten Werkzeuge, Services und Workflows frühzeitig planen – Relevanzabschätzungen durch Datenarchivare und Communities
4.4.4	Interessen zwischen Datenproduzenten, Nachnutzern von Daten und Infrastrukturzentren abstimmen – Mitwirkung in einem gemeinsamen Forum
4.6.2	Datenmanagementpläne als Planungsinstrumente verbindlich einfordern und ihren breiten Einsatz im Forschungsprozess stärker vorantreiben – als Teil guter wissenschaftlicher Praxis
4.7.2	Als Einrichtung bei der Entwicklung und Evaluation von Diensten und Werkzeugen für das Forschungsdatenmanagement auf Nutzerbedürfnisse achten und deren Erfüllung ex post prüfen
4.7.3	In allen Organisationen sowohl Information und Aufklärung als auch juristische Beratung zum rechtssicheren Umgang mit Daten verfügbar machen
4.7.5	Als Forschende die Organisation von Forschungsprozessen unter Bedingungen von Digitalität als Kompetenzfeld begreifen
4.9.1	Verbindliche Verhaltensregeln etablieren, um den Datenfluss zwischen Wissenschaft und Wirtschaft zu gestalten – Ziel: Kultur, die Offenheit und Datenverantwortung miteinander verknüpft
4.12.1	Technisch-organisatorische Maßnahmen zur Datensicherheit bei der Entwicklung von Informationsinfrastrukturen stärker fokussieren
4.12.2	Diskurs über Datenschutz um Aspekte der Datensicherheit erweitern, um auf ein entsprechendes Verhalten der Beteiligten hinzuwirken
4.13.1	Sonderanstrengungen im Rahmen des vor Ort Möglichen, um zum Aufbau einer NFDI beizutragen

LITERATUR

- 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 (2014): Zum Urheberrecht. Entfristung von Paragraph 52a UrhG nur ein Zwischenschritt – Allgemeine Bildungs- und Wissenschaftsschranke muss kommen (Pressemitteilung), o.O., online verfügbar unter: http://www.allianzinitiative.de/fileadmin/user_upload/redakteur/pm_wissenschaftsschranke_2014.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.
- Bauer, Bruno et al. (2015): Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung. Report 2015. Version 1.2, Wien, DOI: 10.5281/zenodo.31935.
- 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 (2014): Strategie der Bundesregierung zum Europäischen Forschungsraum (EFR). Leitlinien und nationale Roadmap, Bonn, online verfügbar unter: http://www.bmbf.de/pub/BMBF_Forschungsstrategie.pdf, zuletzt geprüft am: 25.04.2016.
- BMBF – Bundesministerium für Bildung und Forschung (2014): Internationale Kooperation. Aktionsplan, Berlin, online verfügbar unter: https://www.bmbf.de/pub/Aktionsplan_Internationale_Kooperation.pdf, zuletzt geprüft am: 25.04.2016.
- BMBF – Bundesministerium für Bildung und Forschung (2016): Der Nationale Roadmap-Prozess für Forschungsinfrastrukturen. Investitionen für die Forschung von morgen, Bonn, online verfügbar unter: https://www.bmbf.de/pub/Nationaler_Roadmap_Prozess_fuer_Forschungsinfrastrukturen.pdf, zuletzt geprüft am: 25.04.2016.
- BMBF – Bundesministerium für Bildung und Forschung (2016): European Cloud Initiative. Eckpunkte (unveröffentlicht), Berlin.
- Bundesregierung (2014): Digitale Agenda 2014-2017, online verfügbar unter: https://www.digitale-agenda.de/Content/DE/_Anlagen/2014/08/2014-08-20-digitale-agenda.pdf?__blob=publicationFile&v=6, zuletzt geprüft am: 25.04.2016.
- Chan, Leslie et al. (2002): Budapest Open Access Initiative (Webseite), Budapest, online verfügbar unter: <http://www.budapestopenaccessinitiative.org/read>, zuletzt geprüft am: 25.04.2016.
- Council of the European Union (2016): Standpunkt des Rates in erster Lesung im Hinblick auf den Erlass der Verordnung des Europäischen Parlaments und des Rates zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung). 5419/16. Interinstitutionelles Dossier 2012/0011 (COD), Brüssel, online verfügbar unter: <http://data.consilium.europa.eu/doc/document/ST-5419-2016-REV-1/de/pdf>, zuletzt geprüft am: 25.04.2016.

- Dally, Ortwin/Fless, Friederike/Förtsch, Reinhard (2012): Altermumswissenschaften, in: Neuroth, Heike et al. (Hg.): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme, Boizenburg/Göttingen: Hülsbusch/Universitätsverlag Göttingen, S. 161–178, online verfügbar unter: http://nestor.sub.uni-goettingen.de/bestandsaufnahme/kapitel/nestor_bestandsaufnahme_008.pdf, zuletzt geprüft am: 25.04.2016.
- DARIAH-DE (2016): Memorandum zur nachhaltigen Bereitstellung digitaler Forschungsinfrastrukturen für die Geistes- und Kulturwissenschaften in Deutschland (Webseite), Berlin, online verfügbar unter: <http://dhd-blog.org/?p=6559>, zuletzt geprüft am: 25.04.2016.
- Data Citation Synthesis Group (2014): Joint Declaration of Data Citation Principles (Webseite). Hg. v. Maryann Martone, San Diego, online verfügbar unter: <https://www.force11.org/group/joint-declaration-data-citation-principles-final>, zuletzt geprüft am: 25.04.2016.
- DataCite Metadata Working Group (2015): DataCite Metadata Schema for the Publication and Citation of Research Data. Version 3.1, DOI: 10.5438/0010.
- 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.
- Doove, John (2016): Amsterdam Call for Action on Open Science (Webseite), Amsterdam, online verfügbar unter: <https://wiki.surfnet.nl/display/OSCFA/Amsterdam+Call+for+Action+on+Open+Science>, zuletzt geprüft am: 25.04.2016.
- EFI – Expertenkommission Forschung und Innovation (2016): Gutachten 2016. Gutachten zu Forschung, Innovation und technologischer Leistungsfähigkeit Deutschlands, Berlin, online verfügbar unter: http://www.e-fi.de/fileadmin/Gutachten_2016/EFI_Gutachten_2016.pdf, zuletzt geprüft am: 25.04.2016.
- ERAC – European Research Area Committee and Innovation Committee (2016): Opinion on Open Research Data, Brüssel, online verfügbar unter: http://www.earto.eu/fileadmin/content/Website/ERAC_Opinion_on_Open_Research_Data.PDF, zuletzt geprüft am: 25.04.2016.
- European Commission (2010): Legal framework for a European Research Infrastructure Consortium – ERIC. Practical Guidelines, Brüssel, online verfügbar unter: http://ec.europa.eu/research/infrastructures/pdf/eric_en.pdf, zuletzt geprüft am: 25.04.2016.
- European Commission (2013): Assessing the projects on the ESFRI roadmap. A high level expert group report, Brüssel, online verfügbar unter: <https://ec.europa.eu/research/infrastructures/pdf/jd-final-aegreport-23sept13.pdf>, zuletzt geprüft am: 25.04.2016.
- European Commission (2015): European Research Area. Facts and Figures 2014. EUR 26803 EN, Brüssel, DOI: 10.2777/72967.
- European Commission – Directorate-General for Research and Innovation (2016): Guidelines on Data Management in Horizon 2020. Version 2.1, Brüssel, online verfügbar unter: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_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.
- European Commission (2016): European Cloud Initiative – Building a competitive data and knowledge economy in Europe. COM (2016) 178 final. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Brüssel, online verfügbar unter: <http://data.consilium.europa.eu/doc/document/ST-8099-2016-INIT/en/pdf>, zuletzt geprüft am: 25.04.2016.
- Fecher, Benedikt et al. (2015): A Reputation Economy. Results from an Empirical Survey on Academic Data Sharing (Working Paper Series, 246). Hg. v. RatSWD – German Data Forum, Berlin/Kiel, online verfügbar unter: http://www.ratswd.de/dl/RatSWD_WP_246.pdf, 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.
- G8 (2013): Open Data Charter. Policy Paper, Lough Erne, online verfügbar unter: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/207772/Open_Data_Charter.pdf, zuletzt geprüft am: 25.04.2016.
- GSF – Global Science Forum OECD (2014): International Distributed Research Infrastructures (IDRIS). Issues and Options. Hg. v. OECD – Organisation for economic co-operation and development, Paris, online verfügbar unter: <http://www.oecd.org/sti/sci-tech/international-distributed-research-infrastructures.pdf>, zuletzt geprüft am: 25.04.2016.
- GWK – Gemeinsame Wissenschaftskonferenz: Ergebnisse der Sitzung der Gemeinsamen Wissenschaftskonferenz (GWK) am 28. Juni 2013. Pressemitteilung 11/2013, Berlin/Bonn, online verfügbar unter: <http://www.gwk-bonn.de/fileadmin/Pressemitteilungen/pm2013-11.pdf>, zuletzt geprüft am: 25.04.2016.
- GWK – Gemeinsame Wissenschaftskonferenz (2013): Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen. Ergänztter Bericht des Ausschusses (GWK Drucksache 13.48, unveröffentlicht), Bonn.
- Hanraths, Tobias (2015): Hacker und Missionare. Neue Berufsbilder, in: Süddeutsche Zeitung, online verfügbar unter: <http://www.sueddeutsche.de/karriere/2.220/neue-berufsbilder-hacker-und-missionare-1.2733698>, 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 (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.

- 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.
- 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.
- Klessmann, Jens et al. (2012): Open Government Data Deutschland. Eine Studie zu Open Government in Deutschland im Auftrag des Bundesministerium des Innern, Berlin, online verfügbar unter: http://www.bmi.bund.de/SharedDocs/Downloads/DE/Themen/OED_Verwaltung/ModerneVerwaltung/opengovernment.pdf?__blob=publicationFile, zuletzt geprüft am: 25.04.2016.
- LERU – League of European Research Universities – Research Data Working Group (2013): LERU Roadmap for Research Data. Advice Paper No. 14, o.O., online verfügbar unter: http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf, zuletzt geprüft am: 25.04.2016.
- Meyer-Doerpinghaus, Ulrich/Neuroth, Heike (2015): Die Stärkung von Informationskompetenz im Kontext des Forschungsdatenmanagements: Eine Herausforderung für Hochschulen und Politik, in: ZfBB – Zeitschrift für Bibliothekswesen und Bibliographie 62, Nr. 2, S. 80–84, DOI: 10.3196/186429501562237.
- Nestor (2015): Digitale Langzeitarchivierung in einer Informationsinfrastruktur für Forschung und Bildung, o.O., online verfügbar unter: <http://files.dnb.de/nestor/berichte/nestor-Positionspapier.pdf>, zuletzt geprüft am: 25.04.2016.
- Neuroth, Heike et al. (Hg.) (2012): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme, Boizenburg/Göttingen: Hülsbusch/Universitätsverlag Göttingen, online verfügbar unter: <http://nbn-resolving.de/urn:nbn:de:0008-2012031401>, zuletzt geprüft am: 25.04.2016.
- OECD – Organisation for economic co-operation and development (2008): OECD Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information [C(2008)36]. OECD Ministerial Meeting on the future of Internet Economy in Seoul (Korea) 17-18 June 2008, Paris, online verfügbar unter: <http://www.oecd.org/sti/44384673.pdf>, zuletzt geprüft am: 25.04.2016.
- OECD – Organisation for economic co-operation and development (2015): Making Open Science a Reality. Final Report, Paris, online verfügbar unter: https://www.innovationpolicyplatform.org/sites/default/files/DSTI-STP-TIP%282014%299-REV2_0_0_0_0.pdf, zuletzt geprüft am: 25.04.2016.
- RfII – Rat für Informationsinfrastrukturen (2015): Auftakterklärung, Göttingen, DOI: 10.5281/zenodo.50662.

- 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 (2015): Empfehlungen zur Finanzierung des Nationalen Hoch- und Höchstleistungsrechnens in Deutschland. Drs. 4488-15, Stuttgart, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/4488-15.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.

ONLINERESSOURCEN

Zuletzt geprüft am 25.04.2016. Darstellung in alphabetischer Reihenfolge.

Academia.edu – Platform for academics
<https://www.academia.edu>

DANS – Data Archiving and Networked Services
<http://www.dans.knaw.nl>

da|ra – Registrierungsagentur für Sozial- und Wirtschaftsdaten
<http://www.da-ra.de>

DARIAH-DE – Digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften
<https://de.dariah.eu>

DARIAH-EU – Digital Research Infrastructure for the Arts and Humanities
<https://dariah.eu>

DataCite
<https://www.datacite.org>

Data FAIRport Initiative – Find, Access, Interoperate & Re-use Data
<http://www.datafairport.org>

Deutsche Digitale Bibliothek
<https://www.deutsche-digitale-bibliothek.de>

DFG RIsources (RI = Research Infrastructure)
<http://risources.dfg.de>

DINI Zertifikat – Deutsche Initiative für Netzwerkinformation
<https://dini.de/dini-zertifikat>

DOI – Digital Object Identifier System
<https://www.doi.org>

DSA – Data Seal of Approval
<http://datasealofapproval.org/en/community>

ELIXIR – The European life-sciences Infrastructure for Biological Information
<https://www.elixir-europe.org>
<https://www.elixir-europe.org/news/elixir-and-denbi-agree-collaboration-strategy>

EOSC – European Open Science Cloud
<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

ESFRI – European Strategy Forum on Research Infrastructures
<http://www.esfri.eu/about-esfri>
<http://www.esfri.eu/national-roadmaps>

EUDAT – European data infrastructure
<https://www.eudat.eu/what-eudat>

Europa: Bedarf an Fachkräften – Interview mit Barend Mons vom 9.3.2016
<http://primeurmagazine.com/weekly/AE-PR-05-16-58.html>

Europeana – Europeana Foundation
<http://www.europeana.eu>

FAIR Data Principles – for comment
<https://www.force11.org/group/fairgroup/fairprinciples>

GBIF – Global Biodiversity Information Facility
<http://www.gbif.org>

GEOSS – Group on Earth Observations
<http://www.earthobservations.org/geoss.php>

GFBio – German Federation for Biological Data
<http://www.gfbio.org>

GWVG – Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
<https://www.gwdg.de>

Mendeley – Reference manager and academic social network
<https://www.mendeley.com>
<https://data.mendeley.com>

NCDD – Netherlands Coalition for Digital Preservation
<http://www.ncdd.nl/en/about-the-ncdd>

nestor – Deutsches Kompetenznetzwerk zur digitalen Langzeitarchivierung
<http://www.langzeitarchivierung.de>

nestor-Siegel
http://www.langzeitarchivierung.de/Subsites/nestor/DE/nestor-Siegel/siegel_node.html

NoMaD – The Novel Materials Discovery Laboratory
<https://nomad-coe.eu>

OpenAIRE – Open Access Infrastructure for Research in Europe
<https://www.openaire.eu>

Open Government Partnership Netherlands
<http://www.opengovpartnership.org/country/netherlands>

ORCID-Initiative (Open Researcher and Contributor ID)
<http://orcid.org>

Pangaea – Datenrepositorium für die Erd- und Umweltwissenschaften
<https://www.pangaea.de>

PREPARDE Project – Verzeichnis einiger Data Journals
<http://proj.badc.rl.ac.uk/preparde/blog/DataJournalsList>

ResearchGate – Academic social network
<https://www.researchgate.net>

RADAR – Research Data Repository
<https://www.radar-projekt.org>

RatSWD – Rat für Sozial- und Wirtschaftsdaten
<http://www.ratswd.de>
<http://www.ratswd.de/forschungsdaten/fdz>
<http://www.ratswd.de/forschungsdaten/info>

RDA – Research Data Alliance
<https://rd-alliance.org>

RDA-DE – Research Data Alliance Deutschland
<http://www.forschungsdaten.org/index.php/RDA-DE>

Research Data Netherlands
<http://www.researchdata.nl>

re3data.org – Registry of Research Data Repositories
<http://www.re3data.org>

SOEP – Sozio-oekonomisches Panel
http://www.diw.de/de/diw_02.c.221178.de/ueber_uns.html

Wissenschaftsrat – Evaluation des Leibniz-Instituts für Bildungsverläufe (LIfBi)
<http://www.wissenschaftsrat.de/nc/arbeitsbereiche-arbeitsprogramm/evaluation.html#c20161>

A. Begriffsklärungen

Bericht des Redaktionsausschusses Begriffe an den RfII
vom April 2016

INHALTSVERZEICHNIS

1	Einleitung und Zusammenfassung.....	3
2	Bisherige Ergebnisse – Zentrale Begriffe	4
2.1	Big Data	4
2.2	Communities, Fächer, Disziplinen.....	4
2.3	Creative Commons (Lizenzen), Scientific Commons	5
2.4	Daten.....	6
2.5	Datenfusion, Datenföderation	6
2.6	Datenlebenszyklus	7
2.7	Datenkuration, Datenkuratierung.....	8
2.8	Datenqualität	8
2.9	Datenschutz, Datensicherheit	9
2.10	Digitale Spaltung.....	10
2.11	Digitalisierung, Retrodigitalisierung	11
2.12	Dynamische Wissensintegration	12
2.13	Forschungsdaten, Forschungsdatenmanagement	13
2.14	Forschungsformen	14
2.15	Forschungsinfrastrukturen	14
2.16	Informationen	15
2.17	Informationsinfrastrukturen	15
2.18	Informationskompetenz	16
2.19	Integration, Interoperabilität.....	16
2.20	Long Tail Data	17
2.21	Medien.....	18
2.22	Metadaten	18
2.23	Open Access, Open Data – Data Sharing, Open Source.....	19
2.24	Standards, Standardisierung.....	20
2.25	Zugänglichkeit, Verfügbarkeit	21
3	Empfehlungen – Begriffsverwendung	22
	Literatur und Onlineresourcen	23
	Index.....	29

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 Gestalt 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 niedrighschwelligem 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 urheberrechtlich freien 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 überprüfaren 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 Einzelforschern, 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 Generie-

rung, 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 ‚Wissenslücke‘ (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 (RfII), 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 RfII 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 unbekannt) 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 → Informationsinfrastrukturen
- 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 geschätzt. 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, niedrighschwelliger 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/Identifizierungs-codes über die Metadaten bis zur differenzierten maschinellen Indexierung und/oder intellektuellen Tiefener-schließung 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 langfristigen, standortunabhängigen sowie rechtlich, finanziell und technisch gesicherten 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 EntschlieÙung 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://dublin-core.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://data-sealofapproval.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. Ergänztter 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/whatwe-do/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/archivesite/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_fspres, 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

ANHANG A

Standards8, 12, 13, 15, 16, 17, 18, 19, 20, 21
Verfügbarkeit4, 5, 8, 9, 10, 12, 15, 16, 18, 19, 20, 21, 23
Zugänglichkeit4, 5, 8, 10, 11, 12, 15, 16, 17, 18, 20, 21

B. Die Entwicklung von Konzepten für Informationsinfrastrukturen in der Bundesrepublik Deutschland seit den 1960er Jahren

Kurzfassung eines Berichts des
Redaktionsausschusses Konzepte an den RfII
vom November 2015

INHALTSVERZEICHNIS

1	Zielsetzung und zentrale Ergebnisse	3
2	Analyse der Konzepte	4
2.1	Fachinformationsprogramme und ihre Nachfolger	4
2.2	Nationale und internationale nichtstaatliche Konzepte ab 2000	7
2.3	Neue Strukturdiskussionen in Deutschland 2006 bis 2012	9
2.4	The Rising Tide: Akteursperspektiven seit 2012	12
3	Ergebnisse der Konzeptanalyse	19
	Literatur und Onlineressourcen.....	21

Abkürzungsverzeichnis

BLK	Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung
DFG	Deutsche Forschungsgemeinschaft
GWK	Gemeinsame Wissenschaftskonferenz
IKT	Informations- und Kommunikationstechniken
IuD	Information und Dokumentation
KE	Knowledge Exchange
KfR	Kommission für Rechenanlagen/für IT-Infrastruktur
KII	Kommission „Zukunft der Informationsinfrastruktur“
PSI	Public Sector Information
RDA	Research Data Alliance

1 ZIELSETZUNG UND ZENTRALE ERGEBNISSE

Ausgangspunkt der Arbeit des Redaktionsausschusses Konzepte im Rfll ist die Analyse einer Sammlung von rund 100 nationalen und internationalen Programmen, Studien und Stellungnahmen aus den Jahren 1964 bis 2015. Grundlage der Analyse sind (1) einzelwissenschaftliche Studien, (2) Konzepte und Empfehlungen von Wissenschaftsorganisationen oder Fachgemeinschaften, die unabhängig oder in staatlichem Auftrag entstanden sind, sowie (3) staatliche Programme, die dem konkreten Umsetzen von Konzepten dienen. Schwerpunktmäßig wurden vor allem die Publikationen der beiden letztgenannten Akteursgruppen berücksichtigt.

Um die aktuellen Diskussionen besser einordnen zu können, wurde für die Zeit bis 2000 der Schwerpunkt auf die Fachinformationsprogramme der Bundesrepublik Deutschland mit ihren Auswirkungen auf Strukturentwicklungen und Richtungsentscheidungen gelegt. Nach 2000 wurde das Spektrum der berücksichtigten Veröffentlichungen sehr viel weiter gefasst und auch Akzente in der internationalen Entwicklung aufgenommen.

Da sich das Thema Forschungsdaten zwischen 2000 und 2007 im Zusammenhang mit der Diskussion über den offenen Zugang zu wissenschaftlichen Publikationen zu einem Kernthema der Debatten über Informations- und Forschungsinfrastrukturen entwickelte, wurde bei der Sichtung der aktuellen Veröffentlichungen ein Schwerpunkt auf diesen Bereich gelegt. Aufgrund der exponentiell gestiegenen Zahl von Publikationen seit 2011 erwies sich dieses Vorgehen als notwendig. Für die Zeit ab 2011 wurden somit gezielte Recherchen nach Veröffentlichungen durchgeführt, die Problemanalysen und Lösungsansätze zur Koordination des Handlungsfeldes „Informationsinfrastrukturen für Forschungsdaten“ enthalten.

Ein zentrales Ergebnis der Analyse in größerer zeitlicher Tiefe ist die zunehmende Bedeutung von Forschungsdaten¹ als Handlungsfeld. Dabei erfährt der Begriff eine definitorische Ausweitung auf alle Daten, die in unterschiedlichen disziplinären und institutionellen Kontexten aus Forschungsprozessen entstehen bzw. in diesen verarbeitet werden. Forschungsdaten werden dabei an ganz unterschiedlichen Stellen des Datenlebenszyklus zur Herausforderung für das Datenmanagement und ein integrativer Teil der bis in jüngste Zeit als separate Felder betrachteten Handlungserfordernisse, wie z. B. von Langzeitarchivierung, von Open Access, von virtuellen Forschungsumgebungen, des Hostings oder der Lizenzierung.

Im Laufe der vergangenen sechs Jahrzehnte wurden unterschiedliche und teilweise diametral entgegengesetzte Empfehlungen als Lösungen für die sich ändernden Herausforderungen im Bereich der Informationsinfrastruktur formuliert: institutionelle Infrastrukturen vs. Projektförderung, *top down* vs. *bottom up*, staatlich vs. privatwirtschaftlich, disziplinspezifisch vs. generalistisch. Seit 2011 hat sich die Zahl der Empfehlungen, aber auch die beinahe zeitgleiche Formulierung gegensätzlicher Lösungsansätze deutlich erhöht.

¹ Der Rfll verwendet einen weit gefassten Forschungsdatenbegriff, der sowohl analoge als auch digitale Sammlungen umfasst, vgl. hierzu die Begriffsklärung in Anhang A des Positionspapiers.

2 ANALYSE DER KONZEPTE

2.1 FACHINFORMATIONSPROGRAMME UND IHRE NACHFOLGER

1960er und 1970er Jahre – Der planende Staat

Die Publikation grundlegender Programme und Stellungnahmen zum Thema Forschungsdaten setzt in der Bundesrepublik in den frühen 1960er Jahren mit einer Fokussierung auf den Bereich *Information und Dokumentation* (IuD) und parallel zur Gründung erster spezialisierter Institutionen ein. Die Diskussionen stehen im Zusammenhang mit dem grundlegenden Ausbau des Wissenschafts- und insbesondere des Hochschulsystems infolge des Sputnik-Schocks. In den zunächst noch das Feld beschreibenden Publikationen werden grundsätzlich ein Ansteigen der Informationsangebote und eine zunehmende Ausdifferenzierung der Informationsbedürfnisse in Wissenschaft und Gesellschaft konstatiert. Diese Beobachtungen führen in den 1970er Jahren zu einer qualitativ neuen Phase staatlicher Planung und Förderung des Informations- und Dokumentationswesens, die in entsprechenden Programmen der Bundesregierung gefasst wurden.

Geprägt wird die Diskussion auf staatlicher Ebene durch eine Schrift des Bundesrechnungshofes aus dem Jahre 1962 sowie durch die Einrichtung eines Referates „Wissenschaftliche Dokumentation und Information“ am „Bundesministerium für wissenschaftliche Forschung“ im Jahr 1963.² In dieser Phase wird die flächendeckende Versorgung der Bürger mit Fachinformation als staatliche Aufgabe definiert.

Mit dem PROGRAMM DER BUNDESREGIERUNG ZUR FÖRDERUNG DER INFORMATION UND DOKUMENTATION 1974 BIS 1977³ (IuD-Programm) wird ein breit angelegtes strukturpolitisches Konzept festgelegt. Ausgangspunkt ist die Feststellung, dass Fortschritte in der Forschung und in der technischen Entwicklung, in Wirtschaft, Verwaltung und Politik maßgeblich davon abhängen, ob das in der Welt vorhandene Wissen rasch, konzentriert und aufgabenbezogen zugänglich gemacht und in die Praxis umgesetzt werden kann. Der Begriff der Fachinformation umfasst dabei analog und elektronisch vorliegende Informationen und beinhaltet vor allem Literatur und Dokumentationsstellen, die wissenschaftliche Literatur, Patente und andere Publikationen auswerten. Die Notwendigkeit einer elektronischen Erschließung und Verarbeitung analog vorliegender oder zunächst noch analog anfallender Informationen wird hervorgehoben, um die steigende Menge der Informationen erfassen und verarbeiten zu können.

Das strukturpolitische Konzept des Programmes der Bundesregierung besteht in diesen Jahren im Kern darin, dass 20 inhaltliche Felder der IuD definiert werden. Diese 20 Fachinformationsbereiche sollen in Form von *Fachinformationszentren* – spezialisiert und an verschiedenen Standorten – eingerichtet werden. Dies geschieht in den folgenden Jahren durch Neugründungen und Umwidmungen bestehender Einrichtungen. Einige Fachinformationszentren werden allerdings nie eingerichtet.

² Lechmann (1964) – Dokumentation und Information.

³ BMFT (1975) – IuD-Programm 1974-1977.

1980er bis Mitte 1990er Jahre – Der schlanke Staat

Das Strukturkonzept der 1970er Jahre wurde in den frühen 1980er Jahren den finanziellen Bedingungen, den raschen Entwicklungen der Informationstechnologie und den Veränderungen des Fachinformationsmarktes angepasst und ist durch Standortdiskussionen der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK) modifiziert worden.⁴ Beginnend mit einem Kabinettsbeschluss vom 23. Januar 1980 kommt es zu einem grundlegenden Richtungswechsel, für den das GUTACHTEN ÜBER DIE FACHINFORMATION IN DER BUNDESREPUBLIK DEUTSCHLAND (1983) von besonderer Bedeutung ist.⁵ Es empfiehlt eine Neuformulierung der Fachinformationspolitik, um Förderschwerpunkte festzulegen, die Finanzierung über die Nachfrage zu steuern und die Frage zu klären, „ob es sich bei der Fachinformation um eine Infrastruktur handelt, die der Staat vorzuhalten hat, oder um einen Markt, bei dem der Staat die Rahmenbedingungen setzt“⁶ und damit auch, welche Bereiche staatlicher und welche privatwirtschaftlicher Steuerung überlassen werden sollten. Zur weiteren Beurteilung der Fachinformation nimmt der Wissenschaftsrat 1984 eine Begutachtung vor und berichtet, von den geplanten 20 Einrichtungen seien nur fünf als Informationszentren etabliert worden. Fünf weitere Einrichtungen nähmen zentrale Aufgaben der Fachinformation wahr, ohne den Status von Fachinformationszentren zu haben, so beispielsweise das Deutsche Institut für Medizinische Dokumentation und Information in Köln (DIMDI).⁷

Grundlage für die neue Ausrichtung des Feldes der Fachinformation ist schließlich das FACHINFORMATIONSPROGRAMM 1985-1988 DER BUNDESREGIERUNG.⁸ Der Staat zieht sich mit dem Richtungswechsel programmatisch aus vielen Bereichen der Förderung zurück und definiert weite Teile des Feldes von Information und Dokumentation als private Aufgabe. Es setzt ein Rückbau der Fachinformationszentren ein. Diese Veränderungen finden unter dem Vorzeichen wissenschafts- und hochschulpolitischer Reformen statt, die auch durch eine gleichzeitige Verminderung des Mittelzuflusses und dann einer Mittelstagnation gekennzeichnet ist.

Es kommt zu einer Marktorientierung der Programme und einer Ausrichtung auf den Online-Markt. Fachinformation wird als zentraler Rohstoff definiert. Die Definition umfasst weiterhin analoge und elektronische Informationen, wobei bereits nachdrücklich auf die Zunahme und Notwendigkeit integrierender Lösungen und das Ansteigen an Fachinformationen hingewiesen wird. Es erfolgt zudem eine Neuorientierung nach den verfassungsrechtlichen Zuständigkeiten. Die Ziele des Programms liegen in der Deregulierung und Privatisierung des Fachinformationsmarktes, der Erhöhung des Informationstransfers in der Wissenschaft sowie zwischen Wissenschaft, Wirtschaft und Gesellschaft. Weiteres Ziel ist die Erleichterung des Zugangs zum internationalen Fachinformationsmarkt.

⁴ BMFT (1986) – Fachinformationsprogramm, S. 95.

⁵ Bundesarchiv (Hg.) (1980) – Künftige Förderung der Fachinformationssysteme (Webseite); Bundesrechnungshof (1983) – Gutachten über die Fachinformation.

⁶ BMFT (1986) – Fachinformationsprogramm, S. 95.

⁷ WR (1984) – Stellungnahme zur GID.

⁸ BMFT (1986) – Fachinformationsprogramm.

1996 bis 2005 – Der kooperative Staat

Der politische Richtungswechsel der 1980er Jahre und die grundsätzliche wirtschaftliche Ausrichtung sowie eine Beschränkung auf wissenschaftlich-technische Informationen prägen auch die Programme in den 1990er Jahren. Das FACHINFORMATIONSPROGRAMM 1990-1994 betont besonders die Forderung nach einer weitgehenden Eigenfinanzierung der Fachinformationszentren und belegt dies durch umfangreiche Daten zum stetig wachsenden Kostendeckungsgrad der geförderten Fachinformationseinrichtungen.⁹ Das staatliche Handeln wird hier ganz gezielt auf wissenschaftlich-technische Information (WTI) beschränkt, während die Bereiche der Wirtschaftsinformation und der ressortspezifischen Fachinformation stärker am Markt ausgerichtet werden sollten.

Das Programm für die Jahre 1996 bis 2000 stellt dann die INFORMATION ALS ROHSTOFF FÜR INNOVATION¹⁰ in den Mittelpunkt und definiert Informationen in den Bereichen Wissenschaft, Wirtschaft und Staat als deren Grundlage. Maßgebend ist die Innovationsstimulierung der deutschen Wirtschaft durch wissenschaftlich-technische Information. Zudem tritt hier erstmals „Innovation“ als Ziel programmatisch in den Titel eines Regierungsprogrammes.

Neben der Finanzierung der noch existierenden Fachinformationseinrichtungen ist über die Projektfinanzierung eine programmatische Abkehr von langfristigen Strukturprogrammen zu beobachten. Das Programm stellt das Verhältnis von digitaler und analoger Information sowie den an der Produktion und Bereitstellung beteiligten Akteursgruppen in den Mittelpunkt. Dieses Thema wird mit den Open-Access-Initiativen ab 2000 prägend für die Diskussion über den Zugang zu wissenschaftlicher Information. Gleichzeitig wird mit den Fragen nach der Wirtschaftlichkeit und der staatlichen beziehungsweise privaten Zuständigkeit der Richtungswechsel von den 1970er zu den 1980er Jahren noch einmal explizit angesprochen – ein Spannungsfeld, das bis heute die Diskussionen prägt.

Mit dem Aktionsprogramm der Bundesregierung INNOVATION UND ARBEITSPLÄTZE IN DER INFORMATIONSGESELLSCHAFT DES 21. JAHRHUNDERTS 2000-2005¹¹ wird der Akzent auf die Bedeutung der Informationswirtschaft gelegt. Informations- und Kommunikationstechniken (IKT) rücken in den Vordergrund. So schrieb das BMBF beispielsweise schon 1999 für die Gesundheitsforschung eine IT-Integrationsplattform für die neu gegründeten Kompetenznetze in der Medizin aus. Hieraus entstand die Telematikplattform für Medizinische Forschungsnetze e.V. (TMF) – heute Technologie- und Methodenplattform für die vernetzte medizinische Forschung –, die bis heute den Erfahrungsaustausch in der medizinischen Verbundforschung und einheitliche übergreifende Infrastrukturlösungen befördert.¹²

Insgesamt betrachtet reicht das Spektrum der Handlungsfelder von der breiten Internetanbindung über die Ausbildung bis hin zur Entwicklung innovativer Informations- und Kommunikationstechniken. Die Forschung wird in diesem auf IKT ausgerichteten Aktionsprogramm in Form des technischen Ausbaus des Deutschen Forschungsnetzes (DFN) berücksichtigt.

⁹ BMFT (1990) – Fachinformationsprogramm.

¹⁰ BMBF (1996) – Rohstoff für Innovation.

¹¹ BMBF/BMWi (1999) – Innovation und Arbeitsplätze.

¹² Vgl. BMBF – Kompetenznetze in der Medizin (Webseite).

2.2 NATIONALE UND INTERNATIONALE NICHTSTAATLICHE KONZEPTE AB 2000

2000 bis 2003 – Der freie und umfassende Zugang zum Wissen

Während die frühen Grundsatzpapiere der Bundesregierung bis 2000 dem grundlegenden Aufbau und Ausbau von Fachinformationssystemen¹³ und dann dem Ausbau der technischen Infrastruktur gewidmet waren, beginnt zwischen 2000 und 2003 aus wissenschaftlicher Perspektive eine Serie von öffentlichen Positionierungen und Erklärungen zu grundsätzlichen Fragen des Zuganges zu Informationen im World Wide Web.¹⁴

Am Beginn dieser Phase legt der Wissenschaftsrat im Jahre 2000 seine THESEN ZUR KÜNFTIGEN ENTWICKLUNG DES WISSENSCHAFTSSYSTEMS IN DEUTSCHLAND¹⁵ vor und nimmt auch zur „Nutzung von Informations- und Kommunikationstechnologien“ Stellung. Dabei wird aus der Beobachtung einer durch die modernen Informations- und Kommunikationstechnologien ermöglichten arbeitsteiligen, ortsungebundenen und international organisierten Forschung die Forderung nach der notwendigen technischen Ausstattung und einem möglichst breiten und kostengünstigen Zugang zu allen weltweit elektronisch verfügbaren Wissensressourcen formuliert. Gefordert wird aber auch eine Ermöglichung der Vermarktung des Wissens (Wissensverwertungsgesellschaften) durch die Hochschulen sowie entsprechende Ausbildungsprogramme.

Im Unterschied zur frühen Fachinformationspolitik formuliert der Wissenschaftsrat zudem eine grundlegend veränderte Haltung durch ein umfassendes Internationalisierungsparadigma. Darüber hinaus nimmt er bereits strukturelle Rollenzuweisungen zur intensiveren Nutzung von Informations- und Kommunikationstechnologien vor: Wissenschaftseinrichtungen – insbesondere Hochschulen – müssten sich zur „Avantgarde beim Einsatz digitaler Medien“¹⁶ entwickeln und ihre virtuelle Sichtbarkeit erhöhen.

Die DFG-Kommission für Rechenanlagen (KfR)¹⁷ forderte erstmals 2001 in ihrer Empfehlung INFORMATIVVERARBEITUNG AN HOCHSCHULEN – NETZE, RECHNER UND ORGANISATION von den Universitäten ein standortweites integriertes Informationsmanagement, das sich an einer nachhaltigen Strategie ausrichten sollte. Im Jahr 2016 ging die KfR einen Schritt weiter und empfahl in ihrem Strategiepapier INFORMATIVVERARBEITUNG AN HOCHSCHULEN – ORGANISATION, DIENSTE UND SYSTEME bis 2020 den zunehmenden Aufbau standortübergreifender Lösungen. Die Umsetzung der Empfehlung von 2001 erweist sich bis heute als schwierig, da geeignete Finanzierungsoptionen fehlten und deshalb *bottom up* entstehende Infrastrukturen nach wie vor und oft schlecht koordiniert nebeneinander bestehen.¹⁸

¹³ Thema war dies im Prinzip auch schon in dem Programm für die Jahre 1996 bis 2000 „Information als Rohstoff für Innovation“.

¹⁴ European Council (2001) – Grundsätze von Lund.

¹⁵ WR (2000) – Thesen zur Entwicklung des Wissenschaftssystems.

¹⁶ Ebd., S. 5.

¹⁷ Die Kommission – seit 2011 Kommission für IT-Infrastruktur – veröffentlicht alle fünf Jahre Empfehlungen.

¹⁸ DFG (2001) – Empfehlungen der KfR 2001-2005; DFG (2016) – Empfehlungen der KfR 2011-2015.

Im Rahmen der Forderung einer Diskussion um einen möglichst breiten und kostengünstigen Zugang zu allen weltweit elektronisch verfügbaren Wissensressourcen kommt es national und international zu einer Reihe von Grundsatzserklärungen über den Zugang zum Wissen im Netz. Grundlegende Stellungnahmen sind die BUDAPEST OPEN ACCESS INITIATIVE (14. Februar 2002), das BETHESDA STATEMENT ON OPEN ACCESS PUBLISHING (11. April 2003) und die BERLINER ERKLÄRUNG (22. Oktober 2003)¹⁹, aber auch EMPFEHLUNGEN ZUR DIGITALEN INFORMATIONSVERSORGUNG DURCH HOCHSCHULBIBLIOTHEKEN des Wissenschaftsrates aus dem Jahre 2001.²⁰ Die Digitalisierung vorhandener kultureller und wissenschaftlicher Wissensressourcen und deren Zugänglichkeit in den globalen Netzen sowie der offene Zugang zu aktuellen wissenschaftlichen Ergebnissen über das Internet sind die zentralen Themen, die mit diesen Stellungnahmen angesprochen werden.

Ab 2007 – Forschungsdaten als Themenschwerpunkt

Die Debatte um den freien Zugang zu den Ergebnissen wissenschaftlicher Arbeit und die Forderung nach Open-Access-Publikationen haben in Kombination mit dem grundsätzlich konstatierten Zusammenhang zwischen Information und Innovation sowie dem technischen Ausbau der Informationsinfrastrukturen ein weiteres Thema in den Mittelpunkt gerückt: die *Forschungsdaten*.

Die OECD veröffentlichte 2007 Empfehlungen, in denen sie das öffentliche Interesse an der Erschließung und an einem breiten Zugang zu Daten aus öffentlich geförderter Forschung betont. Begründet wird ein besseres Informationsmanagement u. a. mit effizienteren Investitionen in Forschung und Entwicklung, was letztlich den Mitgliedstaaten nütze.²¹ Der innovativen Forschung wird dabei eine grundlegende und zentrale Rolle zugesprochen, globalen Veränderungen und Herausforderungen zu begegnen, die vom Gesundheitswesen über Klimawandel und erneuerbare Energien bis hin zum Ressourcenmanagement reichen. Die großen globalen Themen machen deutlich, dass man ebenso global Daten zu deren Bewältigung benötigt. Die Mitgliedstaaten werden zur Umsetzung der OECD-Empfehlungen aufgefordert. In späteren Empfehlungen und Berichten finden sich zahlreiche Bezüge auf dieses Papier.²² Die Empfehlungen benennen mit hoher Genauigkeit Prinzipien für das Management von Forschungsdaten, die auch in späteren Empfehlungen und Berichten wiederkehren: darunter Open Data, Transparenz, Qualität und Sicherheit, Nachhaltigkeit, Urheberrecht und Standardisierung.

Auf EU-Ebene werden in den Jahren 2007/2008 zeitgleich in ganz unterschiedlicher Verbindlichkeit zwischen Verordnung, Empfehlung bis hin zur Formulierung einer Digitalen Agenda die Probleme des Zuganges zu Informationen thematisiert. Neben Empfehlungen zu öffentlichen Daten (Public Sector Information – PSI) werden über EU-Verordnungen für Geodaten sogar konkrete Infrastrukturen eingerichtet und definiert.²³

¹⁹ Allianz der Wissenschaftsorganisationen (2003) – Berliner Erklärung; Chan et al. (2002) – Budapest Open Access Initiative (Webseite); Suber et al. (2003) – Bethesda Statement (Webseite).

²⁰ WR (2001) – Empfehlungen zur digitalen Informationsversorgung.

²¹ OECD (2007) – Access to Research Data.

²² Siehe z.B. in: European Commission (2010) – Riding the wave; KII- Kommission Zukunft der Informationsinfrastruktur (2011) – Gesamtkonzept, S. 29.

²³ Die 2007 verabschiedete Richtlinie 2007/2/EG des Europäischen Parlamentes und des Europäischen Rates legt den Grundstein für die Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE).

In unterschiedlichen Veröffentlichungen wird die Bedeutung von Zugang und Verbreitung wissenschaftlicher Information für die Entwicklung der European Research Area (ERA) betont.²⁴

Parallel dazu kommt es zunehmend zu Diskussionen um die Entwicklung der *Information and Communication Technology* (IKT). Programmatisch wird auf europäischer Ebene 2010 EINE DIGITALE AGENDA FÜR EUROPA vorgelegt, in der die Erfolgsbedingungen und Hindernisse zur Mobilisierung des IKT-Potenzials benannt und ein Aktionsplan entworfen wird.²⁵ Diese Ausrichtung war bereits im Programm der Bundesregierung von 2000 bis 2005 sichtbar und spiegelt die generelle Bedeutung der digitalen Wirtschaft innerhalb der Konzeptpapiere wider.

2.3 NEUE STRUKTURDISKUSSIONEN IN DEUTSCHLAND 2006 BIS 2012

„Zukunft der Fachinformation“ (2006) – Generische Diagnose der Problemlage

In ihrem 2006 veröffentlichten Abschlussbericht NEUAUSRICHTUNG DER ÖFFENTLICH GEFÖRDERTEN INFORMATIONSEINRICHTUNGEN²⁶ nimmt die BLK-Arbeitsgruppe „Zukunft der Fachinformation“ die von Bund und Ländern gemeinsam geförderten Fachinformationszentren und Zentralen Fachbibliotheken in den Blick. Diese bilden zusammen mit weiteren Informationseinrichtungen und Bibliotheken ein das gesamte wissenschaftliche Fächerspektrum abdeckendes System der überregionalen Informationsversorgung und gehören der Leibniz-Gemeinschaft an. Das Papier betrachtet also – unter aktuellen Vorzeichen – jene Entwicklungen, die mit dem Regierungsprogramm 1974 einsetzten.

Der BLK-Bericht fordert einen Ausbau der digitalen Angebote zur Entwicklung des Wissenschaftsstandortes Deutschland und führt sehr umfassend aus, welche Entwicklungsschritte Informationseinrichtungen gehen müssen, um digitale Wissenschaft zu ermöglichen. Auch wenn vor allem die Verfügbarkeit von Literatur und Volltexten im Fokus steht, so sind die grundlegenden Diagnosen generisch und finden sich in ähnlicher Form auch in späteren Empfehlungen und Stellungnahmen²⁷. Empfohlen werden die Entwicklung netzbasierter Informationsdienste – vorzugsweise Open Access –, integrativer Angebote für Forschungsdaten und Publikationen, neuer Geschäftsmodelle – angepasst an die neuartigen Rahmenbedingungen digitaler Wissenschaft –, die Bündelung von Ressourcen durch Arbeitsteilung und Kooperationen sowie eine übergreifende strategische Steuerung.

Die Arbeitsgruppe bearbeitet vier Handlungsfelder – Informationsdienstleistungen, Lizenzierung, Open Access und Archivierung – und schlägt erstmalig auch einen Rat für Informationsinfrastrukturen als Koordinationsgremium vor. Der Bericht bezieht sich insbesondere auf Einrichtungen der Leibniz-Gemeinschaft.

²⁴ Council of the European Union (2007) – Schlussfolgerungen zu wissenschaftlichen Informationen; European Commission/European Union (2007) – Richtlinie 2007/2/EG; OECD (2008) – Recommendation Public Sector Information.

²⁵ European Commission (2010) – Digitale Agenda.

²⁶ BLK (2006) – Neuausrichtung Informationseinrichtungen.

²⁷ Siehe z. B.: WR (2012) – Empfehlungen zu Informationsinfrastrukturen; KII- Kommission Zukunft der Informationsinfrastruktur (2011) – Gesamtkonzept; OECD (2007) – Access to Research Data.

Die BLK nimmt die Empfehlungen der Arbeitsgruppe im Oktober 2006 zustimmend zur Kenntnis und beauftragt die Leibniz-Gemeinschaft mit der Erarbeitung eines Strukturkonzeptes für die gesamte Informationsinfrastruktur.

„Zukunft der Informationsinfrastruktur“ (2011) – Differenzierung der Handlungsfelder

Mit der Verlagerung des Bereiches Wissenschaft und Forschung von der BLK an die Gemeinsame Wissenschaftskonferenz (GWK) im Jahr 2008 wird bereits 2009 von der GWK das Thema der Informationsinfrastruktur aufgegriffen. Sie bittet die Leibniz-Gemeinschaft, eine Kommission „Zukunft der Informationsinfrastruktur“ einzusetzen. Mit dieser Entscheidung der GWK, die sowohl die Arbeit der Kommission als auch die Arbeit des Wissenschaftsrats in diesen Jahren prägt, wird eine neue Phase intensiver Diskussionen über die Informationsinfrastruktur in Deutschland ausgelöst, die bis heute ihren Niederschlag in einer stetig steigenden Zahl grundlegender Papiere findet.

Die EMPFEHLUNGEN DER KOMMISSION „ZUKUNFT DER INFORMATIONENINFRASTRUKTUR“ (KII) werden 2011 als „Gesamtkonzept für die Informationsinfrastruktur in Deutschland“ vorgelegt.²⁸ Die wissenschaftliche Infrastruktur wird in diesem Konzept als Bestandteil der Forschungsinfrastruktur mit einem erheblich über die Fachinformation hinausgehenden Aufgabenspektrum definiert. Die ehemals klar abgegrenzten und abgrenzbaren Fach- und Aufgabengebiete hätten, so das Ergebnis der Analyse, an Trennschärfe verloren.

Grundlegend für das KII-Gesamtkonzept ist die Definition des Begriffs *Informationsinfrastruktur* als nationales, disziplinübergreifendes Netz von Einrichtungen. Diese nähmen dezidiert in öffentlichem bzw. institutionellem Auftrag die Versorgung im weitesten Sinne von Wissenschaft und Forschung mit Information und damit zusammenhängenden Dienstleistungen wahr. Vor diesem Hintergrund wird dem Konzept ein ganzheitlicher, strukturorientierter Ansatz zugrunde gelegt. Auf der Grundlage der Analyse werden von der KII acht Handlungsfelder definiert: (1) Lizenzierung, (2) Hosting/Langzeitarchivierung, (3) Nichttextuelle Materialien, (4) Retrodigitalisierung/Kulturelles Erbe, (4) Virtuelle Forschungsumgebungen, (6) Open Access, (7) Forschungsdaten und (8) Informationskompetenz/Ausbildung. Den Handlungsfeldern werden spezifische Einrichtungen und Institutionen zugewiesen, die als Koordinatoren diese Aufgabe für drei bis fünf Jahre wahrnehmen sollen.²⁹

Mit dem KII-Gesamtkonzept ist der Paradigmenwechsel gegenüber den frühen Definitionen von Fachinformation und Fachinformationsinfrastrukturen des IuD-Programms besonders deutlich zu fassen. Der Aufteilung in verschiedene, disziplinär definierte Bereiche von Fachinformationen ist die klare Systematik einer Definition von disziplin-, sparten- und institutionenübergreifenden Handlungsfeldern gegenübergestellt. Die Handlungsfelder sind jedoch parataktisch angeordnet und getrennt. Tatsächlich sind es jedoch inhaltlich eng zusammenhängende und vernetzte Handlungsfelder. Ihre Vernetzung ist auf der Ebene der Einrichtungen angelegt, die die Informationsinfrastruktur in der Definition von KII ausmachen.

²⁸ KII- Kommission Zukunft der Informationsinfrastruktur (2011) – Gesamtkonzept.

²⁹ Die Partnerorganisationen der Allianz der Wissenschaftsorganisationen haben bereits in ihrer Schwerpunktinitiative „Digitale Information“ seit 2008 ähnliche Handlungsfelder identifiziert und die Arbeit an konkreten Lösungen aufgenommen. Vgl. Allianz der Wissenschaftsorganisationen (2008) – Schwerpunktinitiative Digitale Information.

„Weiterentwicklung der Informationsinfrastruktur“ (2011/12) – Einbettung in das selbstorganisierte und wettbewerbsorientierte Wissenschaftssystem

Gleichzeitig mit dem KII-Gesamtkonzept werden im Jahr 2011 vom Wissenschaftsrat ÜBERGREIFENDE EMPFEHLUNGEN ZU INFORMATIONENINFRASTRUKTUREN verabschiedet, die auf Empfehlungen zu FORSCHUNGSINFRASTRUKTUREN IN DEN GEISTES- UND SOZIALWISSENSCHAFTEN, zu WISSENSCHAFTLICHEN SAMMLUNGEN ALS FORSCHUNGSINFRASTRUKTUREN sowie zur ZUKUNFT DES BIBLIOTHEKARISCHEN VERBUNDSYSTEMS IN DEUTSCHLAND aufbauen.³⁰ Der Wissenschaftsrat hält es für erforderlich, darauf aufbauend in einem zweiten Schritt eine Gesamtstrategie für die Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland zu entwerfen und nimmt auf Bitten der GWK auch zum KII-Gesamtkonzept Stellung. Im Jahre 2012 wurden diese EMPFEHLUNGEN ZUR WEITERENTWICKLUNG DER WISSENSCHAFTLICHEN INFORMATIONENINFRASTRUKTUREN IN DEUTSCHLAND BIS 2020³¹ veröffentlicht.

Die Empfehlungen des Wissenschaftsrates verfolgen den Anspruch einer integrativen Weiterentwicklung sowohl der digitalen als auch der nicht-digitalen Infrastrukturbereiche. Der Wissenschaftsrat folgt im Kern der Definition der acht Handlungsfelder des KII-Gesamtkonzepts, erweitert diese und verbindet sie mit einer Typologie von sechs fächerübergreifenden Forschungsformen.³² Dadurch kommt es zu einer stärkeren Einbeziehung unterschiedlicher Typen von Forschungsprozessen, zu einer Integration der Handlungsfelder in den Forschungsprozess und zu einer Berücksichtigung durchaus fachspezifischer Anforderungen an Informationsinfrastrukturen. Der Wissenschaftsrat spricht zusätzliche Empfehlungen zur digitalen Transformation, zu Forschungsdaten sowie zur Langzeitarchivierung und-verfügbarkeit aus. Eine Notwendigkeit zur Erweiterung des KII-Konzeptes sieht der Wissenschaftsrat zudem hinsichtlich der nicht-digitalen Medien und Objekte.

Eindeutig positioniert sich der Wissenschaftsrat hinsichtlich der Grundfrage öffentlicher oder privatwirtschaftlicher Zuständigkeiten. Er betont, dass die Gewährleistung des Zuganges zu den für die wissenschaftliche Arbeit und das Studium (sowie die Bildung insgesamt) erforderlichen Daten, Informationen und Wissensbeständen eine öffentliche Aufgabe ist und bleibt. Die Bearbeitung der drängenden zukünftigen Aufgabenfelder durch fach-, forschungsfeld- und medienbezogene Initiativen sei fortzusetzen und auf wettbewerblicher Basis über ihre Finanzierung sowie über die Finanzierung neuer Initiativen zu entscheiden. Die Weiterentwicklung des Systems der Informationsinfrastrukturen muss aus Sicht des Wissenschaftsrates den unterschiedlichen Funktionen ebenso wie der Heterogenität von Medien als Informationsträgern für Forschung, Studium, Nachwuchsförderung, Technologieentwicklung und Wissenstransfer Rechnung tragen, indem sie sowohl den digitalen als auch den nicht-digitalen Bereich einbezieht. Hierbei sind die Erfordernisse der unterschiedlichen Forschungsformen und Fächerkulturen zu berücksichtigen. Auch damit treten die unterschiedlichen Fachkulturen wiederum in den Fokus.

Anders als die KII empfiehlt der Wissenschaftsrat also, strukturbildende Koordinierungsfunktionen über wettbewerbliche Verfahren zu organisieren und nicht in einem Top-down-Prozess einzelne Forschungs- oder

³⁰ WR (2011) – Übergreifende Empfehlungen zu Informationsinfrastrukturen; WR (2011) – Empfehlungen zu Forschungsinfrastrukturen; WR (2011) – Empfehlungen Wissenschaftliche Sammlungen; WR (2011) – Empfehlungen Bibliothekarische Verbundsystem.

³¹ WR (2012) – Empfehlungen zu Informationsinfrastrukturen.

³² Vgl. Anhang A zum Positionspapier – Begriffsklärungen.

Infrastruktureinrichtungen damit zu beauftragen. Im Kern bekräftigt der Wissenschaftsrat die Notwendigkeit eines übergeordneten Koordinierungs- und Beratungsgremiums zur strategischen Weiterentwicklungen des Gesamtsystems der Informationsinfrastrukturen in Deutschland. Fachgemeinschaften sollen Qualitätskriterien für die Archivierung von Forschungsdaten aufstellen, Informationseinrichtungen sollen stärker untereinander und mit der Wissenschaft kooperieren. Zentrale Aussagen sind darüber hinaus u. a. die Anerkennung der Publikation von Forschungsdaten als eigenständige Forschungsleistung und der Aufbau von spezifischen Kompetenzen durch Lehre und Nachwuchsförderung in diesem Feld.

2.4 THE RISING TIDE: AKTEURSPERSPEKTIVEN SEIT 2012

Forschende – Forschungsdaten im Wissenschaftsprozess

Im Umfeld und in der Folge des 2011 publizierten KII-Gesamtkonzeptes und der Empfehlungen des WR aus den Jahren 2011/2012 ist eine Intensivierung der Diskussion durch unterschiedliche Fachgemeinschaften zu beobachten. Die disziplinär bzw. auch durch transdisziplinäre Forschungsprozesse geprägte Perspektive der Forschenden ist durch eine Fokussierung auf das Thema Forschungsdaten geprägt und hat ein facettenreiches Bild des Umganges mit Forschungsdaten und möglicher Lösungen der zentralen Probleme hervorgebracht. Dies wird hier an drei beispielhaften Veröffentlichungen beleuchtet.

Im Zukunftsreport *LEBENSWISSENSCHAFTEN IM UMBRUCH*³³ der Nationalen Akademie der Wissenschaften Leopoldina wird für die unter „Omics“ zusammengefassten Disziplinen nicht nur ein exponentielles Anwachsen großer und heterogener Datensätze (Big Data) aus Hochdurchsatztechnologien angeführt, sondern es werden auch die damit verbundenen Probleme aufgezeigt. Deutschland sei nicht auf diese Datenmengen vorbereitet. Empfohlen wird der strategische Aufbau einer nationalen Omics- und IT-Infrastruktur als Netzwerk bundesweit verteilter Zentren (Universitäten und Institute), ein massiver Ausbau der IT- und Bioinformatik-Infrastruktur, eine Verknüpfung der Zentren (technisch, Forschung, Ausbildung, Zugang), verbindliche Standards, Qualitätskontrolle (national/international) sowie eine Gestaltung der Ausbildung und der Karrierewege. Die staatliche Verantwortung wird hervorgehoben und eine Finanzierung durch Bundesmittel gefordert.

Wie der Zukunftsreport geht auch die im Rahmen von Nestor, dem Kompetenznetzwerk Langzeitarchivierung, vorgelegte Bestandsaufnahme *LANGZEITARCHIVIERUNG VON FORSCHUNGSDATEN*³⁴ von der konkreten Perspektive disziplinär definierter Akteursgruppen aus. Deutlich tritt in den einzelnen Beiträgen hervor, dass es aufgrund der unterschiedlichen Forschungsprozesse der jeweiligen Akteursgruppen sehr unterschiedliche Definitionen von Forschungsdaten gibt. Die Definitionen und Lösungen (Metadaten) werden als ebenso vielfältig bewertet wie die Forschungsprozesse in den Disziplinen. Unterschiedliche Ansätze bei der Langzeitarchivierung treten dabei jedoch nicht als Ausdruck eines mangelnden Kooperationswillens über die Disziplingrenzen hinweg hervor, sondern als logische Konsequenz der unterschiedlichen Anforderungen und der praktizierten Methoden innerhalb der einzelnen Disziplinen. Betont wurde auch, dass ein kooperativer

³³ Leopoldina (2014) – Lebenswissenschaften im Umbruch (kurz).

³⁴ Neuroth et al. (Hg.) (2012) – Langzeitarchivierung.

Ansatz in den meisten Bereichen, die sich zudem nicht immer streng auf Disziplinen reduzieren lassen, üblich ist. Datenzentren werden von vielen Disziplinen als die ideale Lösung angesehen, jedoch in einer großen Variationsbreite.

Die Variabilität praktizierter Methoden zeigt auch eine Studie der internationalen Initiative Knowledge Exchange (KE) auf: SOWING THE SEED. INCENTIVES AND MOTIVATIONS FOR SHARING RESEARCH DATA, A RESEARCHER'S PERSPECTIVE³⁵. Die KE-Studie untersucht anhand von fünf Fallstudien aus verschiedenen Disziplinen die Praxis des Teilens von Daten anhand definierter Modi. Durch Befragungen von Forschenden wird deutlich, dass die Weitergabe von Daten zur Nachnutzung zwar durchaus üblich ist, jedoch in der Regel selektiv erfolgt bzw. auf einen kontrollierbaren Kreis begrenzt ist. Die KE-Studie kommt zu dem Schluss, dass die individuellen Anreize für das Deponieren von Forschungsdaten in Repositorien noch gering sind und es einer kollektiven Verhaltensänderung bedarf, um das Ideal vernetzter und zugänglicher Forschungsdateninfrastrukturen zu erreichen. In der Konsequenz empfiehlt die Studie, dass Förderer, Forschungseinrichtungen, Fachgesellschaften und Verlage aktiv werden, um Anreize für das Deponieren von Forschungsdaten zu setzen und die Schaffung nachnutzbarer Datensammlungen zu stimulieren.

Wissenschaftsorganisationen – Forschungsdaten in der Wissenschaftspolitik

Die Allianz der deutschen Wissenschaftsorganisationen erarbeitet im Rahmen der Schwerpunktinitiative „Digitale Information“ Vorschläge für verschiedene Handlungsfelder. Als Ergebnis der Arbeitsgruppe „Forschungsdaten“ wird 2014 das Positionspapier RESEARCH DATA AT YOUR FINGERTIPS vorgelegt.³⁶ Darin wird zunächst programmatisch eine Vision bzw. das zu erreichende Ideal im Umgang mit Forschungsdaten an den Anfang gestellt. Wissenschaftlerinnen und Wissenschaftler aller Disziplinen sollen, so die Vision, auf alle Forschungsdaten einfach, schnell und ohne großen Aufwand zugreifen, um auf höchstem Niveau zu forschen und exzellente Ergebnisse zu erzielen. Sie könnten gemeinsam mit anderen arbeiten und ihre Forschungsergebnisse sicher aufbewahren. Forschungsdaten stünden dabei in einer Form zur Verfügung, die Forschung sowohl über disziplinäre als auch über nationale Grenzen hinweg ermöglicht und erleichtert. Die Veröffentlichung von Forschungsdaten und Software steigere die wissenschaftliche Reputation. Wissenschaftlerinnen und Wissenschaftler würden beim Sammeln, Erheben, Erfassen und beim Management ihrer Daten unterstützt. Leicht nutzbare digitale Infrastrukturen sowie wissenschaftliche und technische Informationsspezialistinnen und -spezialisten unterstützten den vollständigen Forschungszyklus.

Die Analyse des Status quo durch die Allianz-Arbeitsgruppe „Forschungsdaten“ fasst in knapper Form die wesentlichen aktuellen Punkte zusammen. Wie in der KE-Studie wird die Bedeutung guter fachspezifischer Kulturen des nachhaltigen Umganges mit Daten hervorgehoben, die einige Fachgemeinschaften bereits entwickelt hätten. Die Definition von Forschungsdaten wird ergänzt um Daten, die in nicht-wissenschaftlichen Kontexten entstehen. Hierzu gehören Daten, die auf gesetzlicher Grundlage oder im Rahmen von Verwaltungsprozessen entstehen (PSI), Daten aus sozialen Netzwerken, dem Gesundheitssektor oder Citizen-Science-Initiativen.

³⁵ KE- Knowledge Exchange (2014) – Sowing the seed.

³⁶ Allianz-Initiative Digitale Information- AG Forschungsdaten (2015) – Research data at your fingertips.

Als problematisch wird u. a. benannt, dass die Finanzierung geeigneter Infrastrukturen für Forschungsdaten immer noch nicht geklärt sei, und Förderung vor allem im Rahmen befristeter Projekte aus nicht immer miteinander abgestimmten Programmen stattfinde, was den Betrieb nachhaltiger Infrastrukturen nicht erlaube. Eine wichtige Voraussetzung für den Umgang mit Forschungsdaten seien begleitende Regelwerke (Policies) und formale Bestimmungen, wie sie etwa gegenwärtig im Pilotprojekt für offene Forschungsdaten in Horizon 2020 mit der Verpflichtung zum Einsatz von Datenmanagementplänen und zur Weitergabe der Forschungsdaten nach Projektende erprobt werden.³⁷

Die Empfehlungen des Positionspapieres richten sich an die politische Ebene und greifen dabei auch ganz unterschiedliche Positionen früherer Diskussionsbeiträge wieder auf. Das grundsätzliche Modell kombiniert damit den seit den 1970er Jahren prägenden disziplinären Ansatz mit der Idee der aus dem KII-Gesamtkonzept bekannten Koordinatoren, die hier jedoch nicht durch Institutionen repräsentiert sind, sondern durch Räte, die verschiedene Kompetenzen bündeln. Gefordert werden aber auch rechtliche Regelungen im Bereich Datenschutz, Urheberrecht und Lizenzierung sowie Standardsetzungen und Policies.

In den Allianz-Organisationen selbst entstehen in dieser Zeit ebenfalls Empfehlungen und Stellungnahmen, in denen institutionelle Rollen reflektiert werden. So veröffentlicht beispielsweise die Deutsche Forschungsgemeinschaft (DFG) 2012 das Positionspapier „DIE DIGITALE TRANSFORMATION WEITER GESTALTEN – DER BEITRAG DER DEUTSCHEN FORSCHUNGSGEMEINSCHAFT ZU EINER INNOVATIVEN INFORMATIONSTRUKTUR FÜR DIE FORSCHUNG“, verfasst vom Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme.³⁸ Es wird konstatiert, dass organisatorische Grundlagen und Forschungsdatenrepositorien in einigen Disziplinen erst entwickelt werden und offene Fragen zu rechtlichen Rahmenbedingungen bestehen. Als spezifische Herausforderung werden die internationale Orientierung und die Nachhaltigkeit der geförderten Projekte genannt. Die DFG appelliert auch an die Trägereinrichtungen in Bund und Ländern, die mit DFG-Mitteln initiierten Maßnahmen und Strukturen zu sichern, da diese letztendlich nur komplementär zum Grundauftrag der geförderten Einrichtungen seien.

Ein weiteres Beispiel ist die von der Hochschulrektorenkonferenz im Mai 2014 verabschiedete Empfehlung MANAGEMENT VON FORSCHUNGSDATEN – EINE ZENTRALE STRATEGISCHE HERAUSFORDERUNG FÜR HOCHSCHULLEITUNGEN. Es wird empfohlen, als zentrale Maßnahmen entsprechende Leitlinien durch die Hochschulen zu formulieren sowie Vereinbarungen mit anderen Hochschulen, außeruniversitären Forschungseinrichtungen und fachspezifischen Dateninfrastrukturen zu unterstützen. Zudem sei es „an den Hochschulleitungen, die Informationskompetenz der Hochschulmitglieder zu stärken und die strukturellen Voraussetzungen für ein effizientes, institutionelles Forschungsdatenmanagement zu schaffen.“³⁹ Bund und Länder sollen den Aufbau von Informationsinfrastrukturen über Ländergrenzen hinweg unterstützen, der Rfll wird zusammen mit der HRK in einer koordinierenden Rolle gesehen.

³⁷ Im April 2016 veröffentlichte die Europäische Kommission eine Stellungnahme, nach der Forschungsdaten künftig „by default“ verfügbar gemacht werden sollen. European Commission (2016) – Data and knowledge economy, S. 6.

³⁸ DFG (2012) – Digitale Transformation.

³⁹ HRK (2014) – Management von Forschungsdaten, S. 1.

Bibliotheken und lokale Informationseinrichtungen – Neue Dienstleistungsaufgaben auf lokaler Ebene

Die Bibliotheken setzen sich 2011/2012 kritisch mit dem KII-Gesamtkonzept und den Empfehlungen des Wissenschaftsrates auseinander und fordern insbesondere eine Berücksichtigung der lokalen Einrichtungen im System der Informationsinfrastrukturen ein. Die AGENDA 2020 DER DEUTSCHEN INITIATIVE NETZWERKINFORMATION⁴⁰ (DINI) benennt das Forschungsdatenmanagement als eine von mehreren neuen Anforderungen an die Dienstleistungen der lokalen Infrastruktureinrichtungen, Bibliotheken, Medien- und Rechenzentren. In einer gemeinsam mit dem Kompetenznetz Langzeitarchivierung gegründeten DINI-nestor-Arbeitsgruppe „Forschungsdaten“ werden Arbeitsschwerpunkte dazu für die kommenden Jahre formuliert und der Ausbau des Wikis „forschungsdaten.org“ mit weiteren Partnern geplant.

Bundesländer – Heterogenität regionaler Digitalisierungsstrategien

Die Bundesländer haben verschiedene Initiativen zu den wissenschaftsnahen Themen der Digitalen Agenda ergriffen. So veröffentlicht die Landesregierung Baden-Württemberg 2014 unter dem Titel E-SCIENCE – WISSENSCHAFT UNTER NEUEN RAHMENBEDINGUNGEN ein Fachkonzept zur Weiterentwicklung der wissenschaftlichen Infrastruktur. Das von einer umfassend besetzten Kommission erarbeitete Dokument enthält Strukturkonzepte für die fünf Handlungsfelder Lizenzierung, Digitalisierung, Open Access, Forschungsdatenmanagement und Virtuelle Forschungsumgebungen. Für jedes Handlungsfeld sind die aktuelle Situation und die Herausforderungen analysiert sowie konkrete Maßnahmen benannt. Das Strukturkonzept referenziert auf wichtige bestehende Empfehlungen und ist insgesamt ein Positivbeispiel für die Koordination von Akteuren und Handlungsfeldern auf Landesebene.

Ein Beispiel im Bereich Open Access ist die 2014 veröffentlichte STRATEGIE 2020 DER LANDESREGIERUNG SCHLESWIG-HOLSTEIN FÜR OPEN ACCESS.⁴¹ Die Landesregierung will hier den offenen Zugang zu den Ergebnissen wissenschaftlicher Forschung und zu den Quellen des kulturellen Erbes fördern und nimmt Bezug auf die Allianz-Initiative „Digitale Information“ sowie auf die Open-Access-Erklärungen (Bethesda, Budapest, Berlin). Die Strategie bleibt pragmatisch und betrachtet zunächst den Bereich Publikationen, sieht aber im Vorwort eine spätere Öffnung in Richtung Forschungsdaten vor.

Der Bericht ACATECH BEGLEITPROZESS ZUR STRATEGIE BAYERN DIGITAL⁴² (2015) gibt Handlungsempfehlungen für eine bayerische Digitalisierungsstrategie und wurde in Kooperation mit dem Münchner Kreis entwickelt. Der aus der Perspektive eines Wirtschaftsstandortes geschriebene Bericht widmet sich vorwiegend Digitalisierungsfragen in Wirtschaft und Gesellschaft. Forschung wird in unterstützender Rolle gefördert (Forschung zu offenen technischen und rechtlichen Fragen, bessere Vernetzung von Wissenschaft und Wirtschaft). Die Bayerische Staatsregierung hat 2015 zur Bündelung bestehender Aktivitäten ein Zentrum Digitalisierung.Bayern (ZD.B) eingerichtet.

⁴⁰ DINI (2015) – Agenda 2020; HRK (2014) – Management von Forschungsdaten.

⁴¹ Landesregierung Schleswig-Holstein (2014) – OA-Strategie 2020.

⁴² acatech/Münchner Kreis (2015) – Begleitprozess Bayern Digital.

Bundesebene – Auf dem Weg zu einer Gesamtstrategie für den digitalen Wandel

Die neuen Strukturdiskussionen prägen auch den SECHSTEN ZWISCHENBERICHT DER ENQUETE-KOMMISSION „INTERNET UND DIGITALE GESELLSCHAFT“* BILDUNG UND FORSCHUNG.⁴³ Schwerpunkt der Analyse sind die Herausforderungen für Bildung und Forschung in der digitalen Gesellschaft und digitale Medien in Forschung und Wissenschaft: Open Access und Open Data. Der Bericht erweitert die Diskussion also, wie andere zeitgleich erscheinende Diskussionsbeiträge, um Forschungsdaten und Forschungsdatenmanagement. Die Handlungsempfehlungen der Enquete-Kommission richten sich insbesondere auf die Handlungsfelder (1) Vernetzung von Datenbanken und Repositorien, (2) Lizenzierung, (3) Forschungsdaten, (4) Nachnutzbarkeit von Digitalisaten und (5) Langzeitarchivierung. Der Bericht reduziert die Herausforderungen von Langzeitarchivierung und Nachnutzbarkeit auf große Portallösungen (Europeana und DDB), die aber nicht geeignet sind, das gesamte Handlungsfeld der Forschungsdaten mit seinen konkreten Problemen der Datenkuratierung, der langfristigen Sicherung und aktiven Nachnutzung komplexer, in Datenbanken vorgehaltener Daten zu regeln.

In der DIGITALEN AGENDA von 2014⁴⁴ sieht die Bundesregierung den digitalen Wandel als eine der zentralen Gestaltungsaufgaben für Wirtschaft, Wissenschaft, Gesellschaft und Politik. Die Politik wird diesen Strukturwandel aktiv begleiten, die Rahmenbedingungen für das Leben, Lernen, Arbeiten und Wirtschaften in der digitalen Welt setzen und allen die Teilhabe am digitalen Wandel ermöglichen. Chancen der Digitalisierung sollen für eine innovative und leistungsstarke Volkswirtschaft und für die Sicherung von Wohlstand und Lebensqualität genutzt werden.

Die Digitale Agenda erfasst das gesamte Spektrum von flächendeckenden infrastrukturellen Voraussetzungen, der Sicherheit des Internets bis hin zur Bedeutung des Netzes: Bürgern die großen Chancen des Lernens, der Fort- und Weiterbildung, des Aufstiegs und der Teilhabe am wirtschaftlichen und gesellschaftlichen Leben zu eröffnen. Es gilt auch, den Austausch von Informationen in der Wissenschaft – von der Grundlagenforschung bis zur Umsetzung der Forschungsergebnisse – zu verbessern und schneller zu innovativen Anwendungen zu finden. Die Zugänglichkeit und Nutzbarkeit von digitalen Informationen sei sicherzustellen.

Zentrale Maßnahmen der Digitalen Agenda zielen auf einen hohen Transfer von Forschungs- und Technologieprogrammen in die Wirtschaft, auf den Ausbau einer Hightech-Strategie als Innovationsstrategie, auf die Erhöhung digitaler Teilhabe, E-Government, E-Learning, auf eine Analyse von Qualifizierungsbedarf und -form, auf Open Access, auf die Förderung der Vernetzung von Forschungsdatenbanken und Virtuellen Forschungsumgebungen sowie auf die Gründung eines öffentlich finanzierten Forschungsinstituts für Internet und Digitalisierung. Zur Umsetzung wird u. a. der nationale IT-Gipfel als Plattform für die Zusammenarbeit von Politik, Wirtschaft, Wissenschaft und Zivilgesellschaft auf die Themen der Digitalen Agenda ausgerichtet.

⁴³ Deutscher Bundestag (2013) – Zwischenbericht Enquete-Kommission.

⁴⁴ Bundesregierung (2014) – Digitale Agenda 2014-2017.

EU-Ebene – Diskurs über Umsetzungsstrategien

Die Intensivierung der Diskussion, wie sie für die Jahre 2011/2012 durch das KII-Gesamtkonzept und die Empfehlungen des WR zu fassen ist, ist zeitgleich auch für die europäische Ebene charakteristisch. In den Jahren 2012 bis 2014 befassen sich allein vier Papiere der Europäischen Kommission mit digitalen Infrastrukturen.⁴⁵ Weitere Diskussionsbeiträge entstehen als Studien und Roadmaps im Umfeld von EU-Projekten und Stakeholder-Gremien.⁴⁶

Zwei Roadmaps konzentrieren sich auf den Zugang zu wissenschaftlichen Informationen. Die Empfehlungen beziehen sich sowohl auf den Zugang zu Publikationen als auch zu Forschungsdaten sowie auf deren Nachnutzung. Ebenso werden die E-Infrastrukturen in den Blick genommen, wobei diese Papiere auch im Kontext des Arbeitsprogrammes 2014 bis 2015 im Rahmen von Horizon 2020 zu sehen sind, in dem die konkreten Ausschreibungen definiert sind, die sich auf Forschungsinfrastrukturen und E-Infrastrukturen beziehen. Beiden Papieren ist gemeinsam, dass sie sehr klar die Herausforderungen und Probleme benennen sowie konkrete Empfehlungen zur Lösung einzelner Felder geben.

Der 2012 publizierte FINAL ROADMAP REPORT GLOBAL RESEARCH DATA INFRASTRUCTURE 2020 ist das Ergebnis eines FP7-Projektes.⁴⁷ Der Bericht beschreibt auf der Grundlage einer systematischen Auswertung internationaler Papiere und Empfehlungen von Organisationen, Projekten, Initiativen und Fördereinrichtungen umfassend und detailliert die Rahmenbedingungen für globale und komplexe Forschungsdateninfrastrukturen. Er gibt u. a. Empfehlungen zur Berücksichtigung sozialer und organisatorischer Aspekte und der Prinzipien von Open Science und Open Data, fordert Standards und Werkzeuge, die Schaffung neuer Berufsbilder und fortschrittlicher Infrastrukturdienstleistungen.

Das „WHITE PAPER“ der e-Infrastructure Reflection Group⁴⁸ von 2013 ist auf das Thema Open Science konzentriert, das Open Access, Open Data und Open Research umschließt. Die internationalen Nutzergruppen, die E-Infrastrukturen benötigen, sollen eine langfristige Strategie für eigene Bedarfe verfolgen; sich an der Schaffung von Informationsinfrastruktur-Dienstleistungen beteiligen und Anreize hierfür setzen sowie zur Entwicklung von Standards beitragen. Doppelentwicklungen im Bereich von Öffentlichkeitsarbeit, Nutzereinbeziehung, Werkzeugen und Services sowie rechtlichen und ökonomischen Rahmenbedingungen sollen vermieden werden.

In ihrem Bericht „THE DATA HARVEST: HOW SHARING RESEARCH DATA CAN YIELD KNOWLEDGE, JOBS AND GROWTH“⁴⁹ formuliert die 2013 gegründete Research Data Alliance (RDA) Argumente für das Teilen von Forschungsdaten: Auf der sozioökonomischen Ebene sind dies Förderung von Wirtschaftswachstum und Schaffung von Jobs, Steigerung der Forschungsproduktivität und Kreativität sowie Stärkung des bürgerschaft-

⁴⁵ European Commission (2012) – Access to Scientific Information; European Commission (2012) – Recommendation on access and preservation; European Commission- DG for Research and Innovation (2013) – Guidelines on Open Access; European Commission (2014) – Horizon 2020 Work Programme 2014.

⁴⁶ Zu nennen sind hier die e-Infrastructure Reflection Group (e-IRG), ESFRI sowie drei Projekte aus dem 7. Forschungsrahmenprogramm, die in jeweils zweijährigen Projektlaufzeiten fundierte Konzepte erarbeitet haben: GRDI2020 Consortium (2012) – Final Roadmap Report; ODE- Opportunities for Data Exchange (2011) – Report on Integration; RECODE (2014) – Policy recommendations.

⁴⁷ GRDI2020 Consortium (2012) – Final Roadmap Report, vgl. FN 46.

⁴⁸ e-IRG (2013) – White Paper (long).

⁴⁹ RDA Europe (2014) – The Data Harvest.

lichen Engagements und bessere Bewältigung von Katastrophen. Es werden die ökonomischen, aber auch sozialen „Kosten des Nicht-Teilens“ benannt und die RDA formuliert konkrete Forderungen an die Europäische Kommission und das Parlament, um Data Sharing zu stimulieren und zu koordinieren: die Entwicklung nationaler Strategien zum Forschungsdatenmanagement, von Anreizen und Fördermodellen, die Förderung internationaler Zusammenarbeit – insbesondere im europäischen Raum, das Vermeiden von Überregulierung durch Einbeziehen wissenschaftlicher Stakeholder und von politischen Richtungswechseln bei erfolgreich angelaufenen Initiativen. Darüber hinaus sei das Augenmerk darauf zu richten, dass das Herstellen von Vertrauen und das Ausbilden von Datenkompetenz wichtige Voraussetzungen für Data Sharing sind.

3 ERGEBNISSE DER KONZEPTANALYSE

Die Analyse kann und will aufgrund der Vielzahl an Publikationen keinen Anspruch auf Vollständigkeit erheben. Es geht vielmehr darum, die großen Linien der Entwicklungen seit den 1960er Jahren aufzuzeigen und gleichzeitig die Veränderungen in der Schwerpunktsetzung herauszuarbeiten.

Ausgehend von der frühen Fokussierung auf Fachinformation und Dokumentation (1960 bis 2005) hat das Thema Forschungsdaten seit 2007 in den Konzeptpapieren deutlich an Bedeutung gewonnen und beherrscht aktuell die politische und wissenschaftliche Diskussion.

Die grundlegenden Handlungsfelder, für die mit Blick auf die Forschungsdaten Lösungen zu finden sind, wurden bereits 2007 im OECD-Papier⁵⁰ benannt. Das exponentielle Ansteigen von Forschungsdaten und das wachsende Bewusstsein über die sich damit eröffnenden Handlungsfelder wurden 2010 programmatisch anhand des Titels des Abschlussberichtes der EU-Expertenkommission RIDING THE WAVE – HOW EUROPE CAN GAIN FROM THE RISING TIDE OF SCIENTIFIC DATA⁵¹ zum Ausdruck gebracht. Die Herausforderung ist hier zugleich als internationale definiert.

Unterschiedliche Auslöser haben in den letzten fünf Jahren auf europäischer und nationaler Ebene zu einer wahren Flut an Papieren geführt, die als Richtlinien, Empfehlungen, Konzepte und Programme von unterschiedlichen Akteuren mit unterschiedlichem Auftrag und unterschiedlichen Intentionen publiziert wurden.⁵² Die Herausforderungen sind dabei in immer größerer Präzision beschrieben worden, während die Vorschläge zur Lösung und Umsetzung deutlich divergieren. Dabei wurden beginnend mit den ersten Regierungsprogrammen in den 1970er Jahren im Laufe der Zeit sehr gegensätzliche Vorschläge gemacht, wie man das Thema in Angriff nehmen soll. Die zugrundeliegenden Ansätze lassen sich in vier Fragen zusammenfassen:

1. Befördert man das Feld durch Einrichtung von Institutionen oder durch wettbewerblich ausgeschriebene Projekte?
2. Sind Top-down-Koordinierungen oder Bottom-up-Prozesse zielführender?
3. Wie viel staatliche Förderung und privatwirtschaftliche Förderung sind notwendig und möglich?
4. Braucht es disziplinspezifische Lösungen oder gibt es übergreifende, allgemein verbindliche Handlungsfelder?

Die Publikationshäufigkeit von Empfehlungen hat sich vor allem in den letzten fünf Jahren so erhöht, dass nahezu zeitgleich gegensätzliche Antworten auf die Fragen vertreten werden. Die Vorschläge zu einer möglichen Beförderung des Feldes lassen sich dabei in mehrere grundsätzliche, strukturelle Ansätze gliedern.

⁵⁰ OECD (2007) – Access to Research Data.

⁵¹ European Commission (2010) – Riding the wave.

⁵² Eine Akteursanalyse hat Ulrich Herb im Auftrag der BBAW 2012 erstellt. Herb (2012) – Gestaltung des wissenschaftlichen Kommunikationssystems.

Mit der Gründung der Fachinformationszentren in den 1970er Jahren wurde zunächst die Vorstellung verbunden, dass man für einzelne disziplinär definierte Gruppen über die Einrichtung von Institutionen die Versorgung der Wissenschaft und Gesellschaft mit Information lösen kann. Ein ganz anderer Zugriff erfolgt beispielsweise über das KII-Gesamtkonzept, das acht disziplinunabhängige, übergreifende Handlungsfelder definiert und diese wiederum existierenden Institutionen als Koordinatoren zuordnet.

Der Wissenschaftsrat übernimmt in seinen Empfehlungen von 2012 die grundsätzliche Konzeption der Handlungsfelder, erweitert diese, da er auch analoge Daten im Blick hat, und fügt ihr eine Typologie der Forschungsformen hinzu. Dadurch erfolgt im Kern ein Wiederanknüpfen an disziplinär definierte Gruppen, was sich auch in der Empfehlung an die wissenschaftliche Gemeinschaft widerspiegelt, eigene Kriterien für die Qualitätssicherung von Forschungsdaten zu entwickeln sowie Forschungsdaten für die Nachnutzung in geeigneten Servicezentren zu veröffentlichen. Aufgrund des handlungsfeldübergreifenden Charakters von Forschungsdaten („integrative Betrachtung“) wird die Aufgabe der Lösungsfindung für die definierten Handlungsfelder somit bei der Fachcommunity und den relevanten Stakeholdern des Wissenschaftssystems verortet. Programmatisch wird in den Empfehlungen des Wissenschaftsrates eine Bearbeitung der Handlungsfelder über wettbewerbliche Verfahren beschrieben. Die wissenschaftlichen Communities fordern hingegen disziplinspezifische Lösungen und Datenzentren. Diese wenigen Beispiele ließen sich durch eine Vielfalt an Vorschlägen erweitern.

Bemerkenswert ist, dass in den Empfehlungen meist generalisierende Lösungen auf Ebene des Wissenschaftssystems vorgeschlagen, nur selten aber konkrete Lösungen für die einzelnen Herausforderungen entwickelt werden. Damit liegen auch nur wenige stringent und umsetzungsorientiert formulierte Empfehlungen vor. Eine Differenzierung zwischen dem Bedarf an dauerhaften Infrastrukturen und der Möglichkeit, Einzelaspekte über eine gezielte Projektförderung zu lösen, wird ebenfalls nicht systematisch vorgenommen, was in der Konsequenz das festgestellte Umsetzungsdefizit der bisherigen Empfehlungen mit zu verursachen scheint. Dadurch ergibt sich für den RfII die Aufgabe, die Skalierungs- und Gegenstandsebenen mit den möglichen Lösungsvorschlägen zu korrelieren.

LITERATUR UND ONLINERESSOURCEN

- acatech- Deutsche Akademie der Technikwissenschaften/Münchner Kreis (2015): acatech Begleitprozess zur Strategie Bayern Digital. Ergebnispapier mit Handlungsempfehlungen, München, online verfügbar unter: http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Publikationen/Sonderpublikationen/acatech_Bayern_Digital_WEB.pdf, zuletzt geprüft am: 25.04.2016.
- Allianz der Wissenschaftsorganisationen (2003): Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen, Berlin, online verfügbar unter: http://openaccess.mpg.de/68053/Berliner_Erklärung_dt_Version_07-2006.pdf, zuletzt geprüft am: 25.04.2016.
- Allianz der Wissenschaftsorganisationen (2008): Schwerpunktinitiative „Digitale Information“ der Allianz-Partnerorganisationen, Berlin, online verfügbar unter: http://www.allianzinitiative.de/fileadmin/user_upload/redakteur/pm_allianz_digitale_information_details_080612.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.
- BLK- Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (2006): Neuausrichtung der öffentlich geförderten Informationseinrichtungen. Abschlussbericht der BLK-Arbeitsgruppe „Zukunft der Fachinformation“ (Materialien zur Bildungsplanung und zur Forschungsförderung, 138), Bonn, online verfügbar unter: <http://www.blk-bonn.de/papers/heft138.pdf>, zuletzt geprüft am: 25.04.2016.
- BMBF- Bundesministerium für Bildung und Forschung/BMWi- Bundesministerium für Wirtschaft und Technologie (1999): Innovation und Arbeitsplätze in der Informationsgesellschaft des 21. Jahrhunderts. Aktionsprogramm der Bundesregierung, Bonn/Berlin.
- BMBF – Bundesministerium für Bildung und Forschung: Kompetenznetze in der Medizin (Webseite), online verfügbar unter: <http://www.gesundheitsforschung-bmbf.de/de/159.php>, zuletzt geprüft am: 25.04.2016.
- BMBF- Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (1996): Information als Rohstoff für Innovation. Programm der Bundesregierung 1996-2000, Bonn.
- BMFT- Bundesministerium für Forschung und Technologie (1975): Programm der Bundesregierung zur Förderung der Information und Dokumentation (IuD-Programm) 1974-1977, Bonn.
- BMFT- Bundesministerium für Forschung und Technologie (1986): Fachinformationsprogramm der Bundesregierung mit Zwischenbilanz 1986, Bonn.
- BMFT- Bundesministerium für Forschung und Technologie (1990): Fachinformationsprogramm der Bundesregierung 1990-1994, Bonn.
- Bundesarchiv (Hg.) (1980): Künftige Förderung der Fachinformationssysteme und Informationseinrichtungen mit besonderer Zweckbestimmung. 162. Kabinettsitzung am Mittwoch, dem 23. Januar 1980. Top 4 (Webseite), Bonn, online verfügbar unter: http://www.bundesarchiv.de/cocoon/barch/0000/k/k1980k/kap1_1/kap2_4/index.html, zuletzt geprüft am: 25.04.2016.
- Bundesrechnungshof (1983): Gutachten über die Fachinformation in der Bundesrepublik Deutschland, Bonn.

Bundesregierung (2014): Digitale Agenda 2014-2017, online verfügbar unter: https://www.digitale-agenda.de/Content/DE/_Anlagen/2014/08/2014-08-20-digitale-agenda.pdf?__blob=publicationFile&v=6, zuletzt geprüft am: 25.04.2016.

Chan, Leslie et al. (2002): Budapest Open Access Initiative (Webseite), Budapest, online verfügbar unter: <http://www.budapestopenaccessinitiative.org/read>, zuletzt geprüft am: 25.04.2016.

Council of the European Union (2007): Ergebnisse der Beratungen des Rates (Wettbewerbsfähigkeit) vom 23. November 2007. Schlussfolgerungen des Rates zu wissenschaftlichen Informationen im digitalen Zeitalter. 15362/07, Brüssel, online verfügbar unter: <http://register.consilium.europa.eu/doc/srv?l=-DE&f=ST%2015362%202007%20INIT>, zuletzt geprüft am: 25.04.2016.

Deutscher Bundestag (2013): Sechster Zwischenbericht der Enquete-Kommission „Internet und digitale Gesellschaft“. Bildung und Forschung. Drucksache 17/12029, Berlin, online verfügbar unter: <http://dipbt.bundestag.de/dip21/btd/17/120/1712029.pdf>, zuletzt geprüft am: 25.04.2016.

DFG- Deutsche Forschungsgemeinschaft (2001): Informationsverarbeitung an Hochschulen – Netze, Rechner und Organisation. Empfehlungen der Kommission für Rechenanlagen 2001-2005, o.O., online verfügbar unter: http://dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/kfr-empfehlung2001_2005.pdf, 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 (2016): Informationsverarbeitung an Hochschulen – Organisation, Dienste und Systeme. Empfehlungen der Kommission für IT-Infrastruktur 2011-2015, Bonn, online verfügbar unter: http://dfg.de/download/pdf/foerderung/programme/wgi/empfehlungen_kfr_2011_2015.pdf, zuletzt geprüft am: 25.04.2016.

DINI- Deutsche Initiative für Netzwerkinformationen e.V. (2015): DINI-Agenda 2020. Infrastrukturen für das wissenschaftliche Arbeiten gestalten, Göttingen, online verfügbar unter: http://dini.de/fileadmin/docs/dini_agenda_2020_juli2015_final.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.

European Commission (2012): Commission Recommendation of 17.7.2012 on access to and preservation of scientific information. C(2012) 4890 final, Brüssel, DOI: 10.2777/975917.

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/European Union (2007): Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE). Fundstelle: Amtsblatt der Europäischen Union, online verfügbar unter: <http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32007L0002&from=EN>, 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): Eine Digitale Agenda für Europa. Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen. KOM(2010)245 endgültig, Brüssel, online verfügbar unter: <http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52010DC0245&from=EN>, 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.
- European Commission (2016): European Cloud Initiative – Building a competitive data and knowledge economy in Europe. COM (2016) 178 final. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Brüssel, online verfügbar unter: <http://data.consilium.europa.eu/doc/document/ST-8099-2016-INIT/en/pdf>, zuletzt geprüft am: 25.04.2016.
- European Council (2001): Die Grundsätze von Lund. Schlussfolgerungen der Expertentagung in Lund, Schweden, 4. April 2001, Lund, online verfügbar unter: ftp://ftp.cordis.europa.eu/pub/ist/docs/digicult/lund_principles-de.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.
- Herb, Ulrich (2012): Empfehlungen, Stellungnahmen, Deklarationen und Aktivitäten wissenschaftspolitischer Akteure zur Gestaltung des wissenschaftlichen Kommunikationssystems. Hg. v. BBAW- Berlin-Brandenburgische Akademie der Wissenschaften, Berlin, online verfügbar unter: http://eprints.rclis.org/22426/1/Expertise_Herb_BBAW_2012_A1b.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.
- 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.

- Landesregierung Schleswig-Holstein (2014): Strategie 2020 für Open Access. Unter Mitarbeit von Klaus Tochtermann, Kiel, online verfügbar unter: http://www.schleswig-holstein.de/DE/Landesregierung/VIII/Presse/PI/PDF/2014/141118_msgwg_OpenAccessStrategie.pdf;jsessionid=0F577D604B2500262A6D-139A1B5089F5?__blob=publicationFile&v=3, zuletzt geprüft am: 25.04.2016.
- Lechmann, Heinz (1964): Dokumentation und Information als Anliegen der Bundesrepublik Deutschland, in: Nachrichten für Information 16, S. 157–166.
- Leopoldina (2014): Lebenswissenschaften im Umbruch – Kurzfassung. Herausforderungen der Omics-Technologien für Deutschlands Infrastrukturen in Forschung und Lehre. Zusammenfassung und Empfehlungen (Zukunftsreport Wissenschaft), Halle a. d. Saale, online verfügbar unter: http://www.leopoldina.org/uploads/tx_leopublication/2014_Zukunftsreport_Kurzfassung_web.pdf, 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.
- ODE- Opportunities for Data Exchange (2011): Report on Integration of Data and Publications, o.O., online verfügbar unter: http://www.stm-assoc.org/2011_12_5_ODE_Report_On_Integration_of_Data_and_Publications.pdf, zuletzt geprüft am: 25.04.2016.
- 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.
- OECD- Organisation for economic co-operation and development (2008): OECD Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information [C(2008)36]. OECD Ministerial Meeting on the future of Internet Economy in Seoul (Korea) 17-18 June 2008, Paris, online verfügbar unter: <http://www.oecd.org/sti/44384673.pdf>, zuletzt geprüft am: 25.04.2016.
- RDA Europe- Research Data Alliance Europe (2014): The Data Harvest. How sharing research data can yield knowledge, jobs and growth. An RDA Europe Report, o.O., online verfügbar unter: <https://rd-alliance.org/sites/default/files/attachment/The%20Data%20Harvest%20Final.pdf>, zuletzt geprüft am: 25.04.2016.
- RECODE- Policy RECommendations for Open Access to Research Data in Europe (2014): Policy recommendations for open access to research data, o.O., online verfügbar unter: http://recodeproject.eu/wp-content/uploads/2015/01/recode_guideline_en_web_version_full_FINAL.pdf, zuletzt geprüft am: 25.04.2016.
- Suber, Peter et al. (2003): Bethesda Statement on Open Access Publishing (Webseite), Chevy Chase, online verfügbar unter: <http://legacy.earlham.edu/~peters/fos/bethesda.htm>, zuletzt geprüft am: 25.04.2016.
- WR- Wissenschaftsrat (1984): Stellungnahme zur Gesellschaft für Information und Dokumentation. Drs. 6726/84, Berlin.
- WR- Wissenschaftsrat (2000): Thesen zur künftigen Entwicklung des Wissenschaftssystems in Deutschland. Drs. 4594/00, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/4594-00.pdf>, zuletzt geprüft am: 25.04.2016.

WR- Wissenschaftsrat (2001): Empfehlungen zur digitalen Informationsversorgung durch Hochschulbibliotheken. Drs. 4935/01, Greifswald, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/4935-01.pdf>, zuletzt geprüft am: 25.04.2016.

WR- Wissenschaftsrat (2011): Empfehlungen zu Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften. Drs. 10465-11, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/10465-11.pdf>, zuletzt geprüft am: 25.04.2016.

WR- Wissenschaftsrat (2011): Empfehlungen zu wissenschaftlichen Sammlungen als Forschungsinfrastrukturen. Drs. 10464-11, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/10464-11.pdf>, zuletzt geprüft am: 25.04.2016.

WR- Wissenschaftsrat (2011): Empfehlungen zur Zukunft des bibliothekarischen Verbundsystems in Deutschland. Drs. 10463-11, Berlin, online verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/10463-11.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.

C. Szenarien des Forschungsdatenmanagements

Aus HRK (2015): „Wie Hochschulleitungen die Entwicklung des Forschungsdatenmanagements steuern können“. Empfehlungen der 19. Mitgliederversammlung der HRK am 10. November 2015 in Kiel.

Die HRK hat 2015, um praxisnah die Heterogenität von Herausforderungen des Forschungsdatenmanagements zu zeigen, verschiedene Typen von Forschungsaktivitäten beschrieben. Der Rfll gibt dies im Folgenden wieder.¹

C.1 Abschlussarbeiten und unabhängige Promotionsprojekte	
Daten	Die Datenmenge und die Datenbeschaffenheit können abhängig von der Disziplin sehr heterogen sein. Während bei Promotionsprojekten das FDM komplex und aufwendig sein kann, ist dies bei studentischen Arbeiten aufgrund der zeitlichen Dauer der Projekte i. d. R. weniger der Fall.
Zeitachse	Formelle Fristen werden durch Prüfungsordnungen vorgegeben. Die für die eigentlichen Ausarbeitungen geltenden Fristen sind auch auf die verwendeten und erzeugten Forschungsdaten anzuwenden. Zumindest ist die Mindestaufbewahrungsdauer zu beachten, innerhalb derer eine Aberkennung des Abschlusses möglich ist. Bei Stipendien ist zumindest eine Verpflichtung der unterstützenden Hochschule analog zu angestellten Doktorandinnen und Doktoranden in Erwägung zu ziehen.
Art der Nachnutzung	Bei Stipendien gelten ähnliche Nachnutzungsbedingungen wie bei übrigen Einzelprojekten. Eine Nachnutzung von studentischen Arbeiten ist oft völlig offen. Es obliegt dabei meist den Lehreinheiten, eine Nachnutzung zu unterstützen während es hochschulweite Vorschriften in der Regel nicht gibt.
Relevanz der Daten	Der Wert kann ein weites Spektrum abdecken und zeigt sich oft erst später.
Institutionen	Die Hochschule.
Kosten/Ressourcen	Während die Kosten im Einzelfall überschaubar sind, können sie in der Summe erheblich sein. Verpflichtungen auf Seiten der Hochschule bestehen lediglich im Rahmen der Prüfungsordnung. Drittmittel stehen in der Regel nicht zur Verfügung.
Beratungsbedarf	Hoch.
Probleme	Einerseits tragen die jeweiligen Verfasserinnen und Verfasser grundsätzlich die volle Verantwortung für die Datenpflege, andererseits können die Hochschulen über die Prüfungsordnungen Rahmenvorgaben definieren, die sich auch auf das Datenmanagement erstrecken. Für die Notwendigkeit des FDM müssen Entscheiderinnen bzw. Entscheider oft erst sensibilisiert werden, da dieser Bereich meist dezentral organisiert ist. Es gibt keine Standards. Der Controlling-Aufwand ist hoch.

¹ HRK (2015) – Forschungsdatenmanagement, S. 27-33 (http://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-10-Publikationsdatenbank/Beitr-2016-01_Forschungsdatenmanagement.pdf; zuletzt geprüft am 25.04.2016).

C.2 Einzelprojekte ohne Einbettung in eine größere organisatorische Struktur

(z. B. Projekte im DFG-Normalverfahren und Projekte, die durch die Grundfinanzierung der Hochschulen getragen werden)

Daten	Die Datenmenge ist unterschiedlich groß und heterogen.
Zeitachse	Die Daten sind auf unbestimmte bzw. auf unterschiedliche Dauer zu archivieren, formelle Fristen existieren allenfalls aufgrund der Regeln der guten wissenschaftlichen Praxis.
Art der Nachnutzung	Die Nachnutzung ist oft völlig offen, da diese Projekte meist erst durch Veröffentlichungen (teilweise aber bereits in Form von Preprints) bekannt werden und damit das Interesse an einer Nachnutzung oft mit erheblicher Verzögerung eintritt.
Relevanz der Daten	Ähnliche Situation wie bei [C.1]. Allerdings sollte die Abschätzung des Wertes bereits Gegenstand der Projektbeschreibung sein.
Institutionen	Für das FDM ist i. d. R. die Hochschule verantwortlich. Von Förderorganisationen können zusätzliche Bedingungen gestellt werden.
Kosten/Ressourcen	Die Kosten sind im Einzelfall oft überschaubar, können in der Summe aber erheblich werden. Die DFG stellt bislang keine zusätzlichen Ressourcen für die Grundstufe des Datenmanagements (reine Archivierung) zur Verfügung. Projektspezifische Aufwendungen für weiterführende Stufen des Datenmanagements, die der Nachnutzbarkeit der Forschungsdaten dienen, können mit den Projekten beantragt und bewilligt werden. ² Die Programmpauschale wird bereits für andere Zwecke benötigt und reicht für alle Gemeinkosten nicht aus. Aufseiten der Hochschule werden Ressourcen für das FDM bisher kaum berücksichtigt.
Beratungsbedarf	Hoch, da bisher keine umfassende Ausbildung in diesem Bereich existiert.
Probleme	Für die Notwendigkeit des FDM müssen Entscheiderinnen und Entscheider meist erst sensibilisiert werden. Für diese Projekte existieren bisher kaum Standards und Controlling-Strukturen. Vereinzelt existieren aber bereits fachspezifische Leitlinien (Biodiversitätswissenschaften, Erziehungswissenschaften, Sozialwissenschaften), die die Anforderungen an das FDM konkretisieren und Bewertungsmaßstäbe für FDM etablieren helfen.

² Dies beinhaltet auch Nutzungsgebühren und finanzielle Aufwendungen, die bei der Nutzung bereits etablierter Datenrepositorien anfallen.

C.3 Projekte im Rahmen einer stark vernetzten internationalen Fachcommunity

(z. B. archäologische Grabungsdaten, soziologische oder wirtschaftswissenschaftliche Arbeit mit Massendaten, Sprachwissenschaft)

Daten	Die Datenmengen sind oft groß bis sehr groß. Die Datenbeschaffenheit ist innerhalb des Projekts oder sogar innerhalb der Community weitgehend homogen.
Zeitachse	Kann unterschiedlich sein, da die Community den Rahmen vorgibt. Teils maximal lang (Archäologie).
Art der Nachnutzung	Die Standardnutzung ist vergleichsweise klar. Eine langfristige Strategie der Nachnutzung existiert in vielen Communities aber bisher noch nicht.
Relevanz der Daten	Oft hoch bis sehr hoch.
Institutionen	Zentrale (teils international abgestimmte) Angebote existieren oder sind im Aufbau oft durch außeruniversitäre Einrichtungen/Datenzentren.
Kosten/Ressourcen	Sehr hoch, drittmittelfähig, aber langfristig unüberschaubar.
Beratungsbedarf	i. d. R. gering, da die Strukturen bereits festgelegt und weitgehend akzeptiert sind.
Probleme	Entscheidungen fallen jenseits der lokalen Handlungsebene.

C.4 Befristete universitäre Verbünde oft mit multi- oder interdisziplinärer Zusammensetzung

(z. B. SFB Transregio, Exzellenzprojekt)

Daten	Der Umfang der Datenmengen und die Datenbeschaffenheit sind heterogen und hängen von den jeweiligen Disziplinen ab.
Zeitachse	Wird zunehmend durch eine Verpflichtung gegenüber dem Drittmittelgeber vorgegeben. Sie orientiert sich oft nur an der Projektdauer, obwohl oft ein längerfristiger Bedarf besteht.
Art der Nachnutzung	Heterogen, aber mit zeitnaher Nachnutzung ist i. d. R. zu rechnen. Eine langfristige Nachnutzung findet zurzeit nur vereinzelt statt.
Relevanz der Daten	Unterschiedlich, sie kann abhängig von der Disziplin aufgrund der hohen Projektdauer und damit des Datenumfanges sehr hoch sein.
Institutionen	Konfligierende Optionen (gemeinsam am Ort? gemeinsame überregionale Lösung? Auf Fachcommunities verteilt?). Die Fokussierung auf einen Ort widerspricht der angestrebten Bedeutung des Projekts im Rahmen der Community.
Kosten/Ressourcen	Einzelfallabhängig, Drittmittel nur wenn mitbeantragt. Eine Beantragung ist möglich (z. B. INF-Projekte der DFG).
Beratungsbedarf	Kann sehr hoch sein, da oft neue Strukturen aufgebaut werden müssen.
Probleme	Nicht immer existieren Planungen, Standortkonkurrenzen bergen Konfliktpotenzial.

C.5 Kollaborative, auf internationaler Ebene von vornherein hoch vernetzte Forschung, oft mit Großgeräten

(z. B. in der Klimaforschung, Teilchenphysik, Bioinformatik, Weltraumforschung)

Daten	Die Datenmengen sind in der Regel sehr groß. Die Datenbeschaffenheit ist innerhalb der Community homogen. Es gibt klare Anforderungen an Metadaten etc.
Zeitachse	Lang. Community gibt Routinen vor.
Art der Nachnutzung	Standardnutzung.
Relevanz der Daten	Hoch bis sehr hoch.
Institutionen	International arbeitsteilig verankerte Angebote existieren.
Kosten/Ressourcen	Sehr hoch, drittmittelfähig, aber langfristig unüberschaubar.
Beratungsbedarf	Gering, da die Rahmenbedingungen vorgegeben und unveränderbar sind.
Probleme	Pfadentscheidungen fallen jenseits der lokalen Handlungsebene, Hochschulen sind i. d. R. oft ein Partner unter mehreren.

C.6 Industriekooperation

(z. B. im Engineering, in der Organisationspsychologie, Betriebswirtschaft, Informatik)

Daten	Datenmengen sind i. d. R. klein bis mittelgroß. Die Datenbeschaffenheit ist unterschiedlich.
Zeitachse	Kurz- oder mittelfristig (Vereinbarungssache).
Art der Nachnutzung	Wissenschaftliche Nutzung und/oder privatwirtschaftliche Verwertung, muss verhandelt werden.
Relevanz der Daten	Wissenschaftlich und wirtschaftlich unterschiedlich.
Institutionen	Hochschule oder Auftraggeber (nach Vereinbarung).
Kosten/Ressourcen	Drittmittelfähig im Rahmen des Auftrags.
Beratungsbedarf	Gegeben (insbesondere bei Kooperation mit KMU), neben technischen Fragen ist oft Rechtsberatung erforderlich.
Probleme	Der Partner stellt oft hohe Anforderungen an Vertraulichkeit (und ggf. Sicherheit) von Daten. Die wissenschaftliche Nachnutzung kann auf Rechtsprobleme stoßen. Für die Hochschule müssen Haftungsrisiken ausgeschlossen werden. Der Controlling-Aufwand ist hoch.

D. Daten und Fakten

Die nachfolgenden Darstellungen ergänzen Ausführungen des Positionspapiers um einige Daten und Fakten. Die Aktivitäten im Bereich des Forschungsdatenmanagements sind überaus heterogen und hier nur ansatzweise erfasst. Wie unter 4.10 ausgeführt, sind weitere Anstrengungen erforderlich, um das Thema im Verlauf der Entwicklung für die wissenschaftliche Selbstverwaltung und die Forschungspolitik transparent zu halten.

D.1 AN FORSCHUNGSDATENREPOSITORIEN BETEILIGTE EINRICHTUNGEN

Als „Momentaufnahme“ existierender Forschungsdatenmanagement-Dienste hat der RfII Angaben zu Datenrepositorien mit deutscher Beteiligung aus dem internationalen Register re3data.org analysiert (www.re3data.org). Der Datensatz stammt vom Dezember 2015 und umfasst 220 Repositorien und die jeweils beteiligten Einrichtungen („institutions“) mit ihren Verantwortlichkeiten.¹ Die Daten basieren auf Selbstangaben der Repositorienbetreiber und einer redaktionellen Prüfung durch das re3data-Team.

Tabellarische Übersicht Partnerschaften und Mittelgeber

Zeile 2 und 3: Für die Analyse der Partnerschaften wurden Repositorien ausgewertet, an denen mehr als eine Einrichtung beteiligt ist. Einrichtungen mit der alleinigen Verantwortlichkeit „funding“ oder „sponsoring“ wurden zuvor über einen entsprechenden Filter ausgeschlossen.

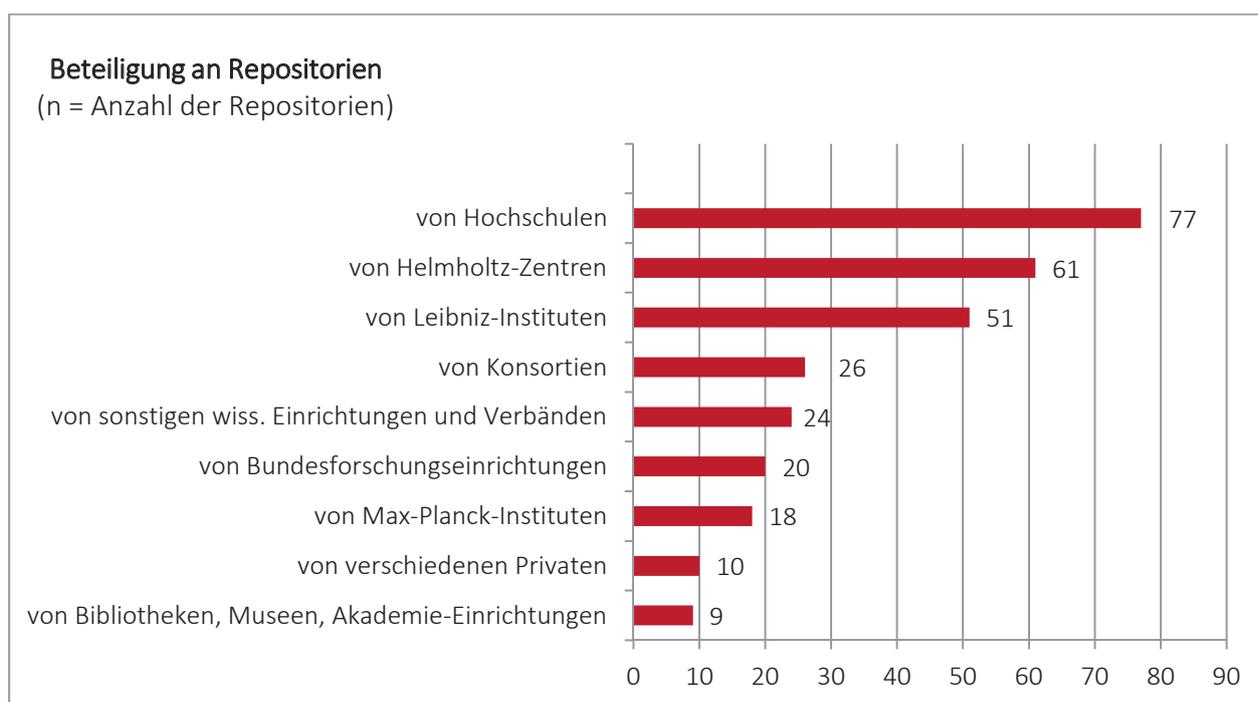
Zeile 4-9: Um die Mittelgeber zu analysieren, wurde der Datensatz nach Einrichtungen mit den Verantwortlichkeiten „funding“ oder „sponsoring“ gefiltert (allein oder in Kombination mit anderen Verantwortlichkeiten). In der Tabelle sind die Ergebnisse für ausgewählte Mittelgeber dargestellt, bei denen aufgrund der Förderbedingungen eine hohe Motivation für vollständige Angaben anzunehmen ist. Darüber hinaus sind vor allem die tragenden Einrichtungen als Förderer oder Sponsoren benannt.

Zeile	Auswertung	Anzahl Repositorien	Anteil
1	Repositorien mit deutscher Beteiligung (gesamt)	220	
2	...davon mit zwei oder mehr Partnern (ohne Mittelgeber)	137	62 %
3	...davon mit mindestens einem internationalen Partner	70	32 %
4	Mittelgeber		
5	...Bundesministerien	41	19 %
6	...DFG	30	14 %
7	...EU	19	9 %
8	...Landesministerien	9	4 %
9	Sonstige	nicht dargestellt	

¹ Pampel et al. (2016): Forschungsdaten-Repositorien mit Beteiligung deutscher Institutionen. Eine re3data.org-Analyse. Version 1.0, <http://dx.doi.org/10.5281/zenodo.50148>.

Auswertung der deutschen Partnereinrichtungen nach Typ

Nachstehend sind die Beteiligungen an den 220 Repositorien im Detail analysiert. Einrichtungen mit der alleinigen Verantwortlichkeit „funding“ oder „sponsoring“ wurden über einen entsprechenden Filter ausgeschlossen. Im Einzelfall vorkommende Doppelnennungen (z. B. zwei Abteilungen innerhalb einer Einrichtung) sind nicht bereinigt. Die Kategorisierung der Partnereinrichtungen aus dem Datensatz von Pampel et al. (2016) wurde an die Bedarfe des Rfll angepasst. Die dargestellten Institutionstypen sollen vor allem die Bandbreite der beteiligten Akteure und verschiedene Sektoren deutlich machen. Die angegebenen Zahlen beziehen sich auf die Repositorien – so sind z. B. Hochschulen an 77 der 220 Repositorien beteiligt.



Legende

Hochschulen	Gruppen, Institute und andere Einrichtungen an Universitäten und Fachhochschulen
Helmholtz-Zentren	Einrichtungen der Helmholtz-Gemeinschaft
Leibniz-Institute	Einrichtungen der Leibniz-Gemeinschaft
„Konsortien“	(Projekt-)Verbünde, die als Partner oder Betreiber für ein Repository angegeben werden (z. B. BIOACID, CLARIN-D): Sie werden in der Analyse als eine Art „kollaborativer Akteur“ sichtbar. Die Beteiligung einzelner Konsortialpartner an dem jeweiligen Dienst wird im Datensatz meist zusätzlich ausgewiesen und ist dann in der Analyse ebenfalls berücksichtigt.
Sonstige wiss. Einrichtungen und Verbände	Eigene Kategorie – Zusammenfassung von Einrichtungen mit jeweils geringeren Beteiligungsanteilen: Landesforschungseinrichtungen, Fraunhofer-Institute, verschiedene wissenschaftliche Verbände und wissenschaftliche Serviceeinrichtungen, die nicht den Wissenschaftsorganisationen zuzuordnen sind.
Bundesforschungseinrichtungen	Hier sind vom Betreiber re3data.org neben den Ressortforschungseinrichtungen auch die aml. Statistik und andere Forschungsdatenzentren aus Bundesbehörden bzw. Bundesunternehmen eingeordnet (z. B. Deutsche Bundesbank).
Max-Planck-Institute	Einrichtungen der Max-Planck-Gesellschaft
Verschiedene Private	Eigene Kategorie – umfasst privatwirtschaftliche Akteure, Stiftungen und durch Stiftungen finanzierte Einrichtungen.
Bibliotheken, Museen, Akademie-Einrichtungen	Universitätsbibliotheken werden als Einrichtungen der Hochschulen behandelt, Forschungsmuseen der Leibniz-Gemeinschaft als Leibniz-Institute.

D.2 EVALUATIONSKRITERIEN FÜR FORSCHUNGSINFRASTRUKTUREN – DEUTSCHLAND UND EU IM VERGLEICH

In Ergänzung zu den Ausführungen unter Abschnitt 2.4 sind hier die Evaluationskriterien der nationalen Roadmap Forschungsinfrastrukturen (BMBF)² und des European Strategy Forum on Research Infrastructures³ gegenübergestellt.

BMBF Roadmap		ESFRI	
Wissenschaftsgeleitete Bewertung	Wirtschaftliche Bewertung	Scientific evaluation	Assessment of maturity
<ul style="list-style-type: none"> ▪ Wissenschaftliches Potenzial ▪ Nutzung ▪ Umsetzbarkeit ▪ Bedeutung für den Wissenschaftsstandort Deutschland 	<ul style="list-style-type: none"> ▪ Finanzierungskonzept (Finanzierungsstruktur, Kosten der Aufbau-, sowie der Nutzungs- und Auslaufphase, wirtschaftliche Risikoabschätzung) ▪ Umsetzungskonzept (Projektpläne, Management-Konzept, Governance, Risikoanalyse) ▪ Nutzungskonzept (Bedarfs- und Zielgruppenanalyse, Zugangsmanagement, Service, Geschäftsplan, Datennutzungs- und Datenmanagement-Konzept) 	<ul style="list-style-type: none"> ▪ scientific impact/ scientific relevance for the respective scientific area in the European RI ecosystem ▪ European added value/pan-European relevance ▪ socio-economic impact ▪ e-needs 	<ul style="list-style-type: none"> ▪ preparatory work achieved ▪ commitment by stakeholders ▪ planning ▪ governance, scientific and legal management ▪ HR policy and project management ▪ costs and financial commitments ▪ feasibility and risks

² 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).

³ ESFRI- European Strategy Forum on Research Infrastructures (2014): 2016 ESFRI Roadmap. Short Guide for Applicants, Brüssel, online verfügbar unter: https://ec.europa.eu/research/infrastructures/pdf/esfri_roadmap_short_guide_21102014.pdf (zuletzt geprüft am: 25.04.2016).

D.3 BEISPIELE FÜR ZENTRALE DIENSTE UND ERMÖGLICHUNGSSTRUKTUREN IN DER NFDI

Diese Beispiele ergänzen die Ausführungen im Abschnitt 4.2.4 und erheben keinen Anspruch auf Vollständigkeit. Zu den deutlich breiter gefassten Aufgabengebieten der NFDI, wie z. B. der Bereitstellung generischer Datenservices und Datenspeicher sowie der Langzeitarchivierung/-verfügbarkeit, siehe Ausführungen unter 4.2.2 und 4.3 des Positionspapiers.

- Zugangsportal mit Zugriffsrechtenmanagement;
- Suchfunktionen, die auch semantische Suchanfragen ermöglichen (Suchbegriffe werden ins Vokabular der Communities übersetzt);
- Dienste zur Registrierung von Daten oder ‚Datenprodukten‘, um das Referenzieren von Daten in Publikationen zu ermöglichen;
- Dienste zur Publikation, dem Austausch bzw. zum kooperativen Bearbeiten von Forschungsdaten, die den einfachen und schnellen Aufbau einer gemeinsamen, ggf. auch geschlossenen Forschungsdatenplattform für unterschiedliche Forschergruppen erlauben;
- Dienste zur Unterstützung der Integration von Daten unterschiedlicher Communities;
- Zugänge zu Datenanalyseressourcen, Visualisierungs- und HPC-Systemen sowie Dienste für den Datentransfer von Repositorien dorthin;
- Technisch-organisatorische Maßnahmen zur Ausfall- und Datensicherheit.

D.4 DRITTMITTELFÖRDERUNG FÜR INFORMATIONSIINFRASTRUKTUREN

Betrachtet wurden Aktivitäten der großen Fördereinrichtungen, die Möglichkeiten zur Förderung digitaler Informationsinfrastrukturen bieten. Gemein ist den nachstehend aufgeführten Programmen das Ziel, die Vernetzung und Bildung von Kooperationsverbänden zu fördern, die über Disziplinen- und Einrichtungsgrenzen hinweg entstehen. So sehen die BMBF-Ausschreibungen die Bildung von Zentren – in der Medizin-informatik den Aufbau von Datenintegrationszentren an Universitätskliniken, in den eHumanities von Kompetenzzentren für die Geisteswissenschaften – vor, in denen Dienste gebündelt werden können. Auch die verschiedenen Programme von BMBF, DFG und Volkswagen-Stiftung zur Forschungsförderung in Museen benennen Vernetzung als Förderziel, das mithilfe der Unterstützung von Veranstaltungen und dem Aufbau von virtuellen Forschungsumgebungen erreicht werden soll.

Die Ausschreibungen der Europäischen Kommission im Horizon2020-Programm, der Trans-Atlantic Platform (T-AP) unter Beteiligung der DFG und die Ausschreibung der Volkswagen-Stiftung für die *Computational Social Sciences* fordern explizit europäische und internationale Forschungsk Kooperationen. Die BMBF-Ausschreibung „Die Sprache der Objekte“ oder die DFG-Ausschreibung „Virtuelle Forschungsumgebungen“ zielen ebenfalls auf die Förderung internationaler Zusammenarbeit. Einige der Förderrichtlinien betonen darüber hinaus die Bedeutung der internationalen Anschlussfähigkeit nationaler Infrastrukturen (DFG – Infrastrukturen für Forschungsdaten, GWK – Forschungsbauten).

Die nationale und internationale Gewinnung qualifizierter Forschender durch Nachwuchsförderung und attraktive Forschungsbedingungen als Voraussetzung für den Kompetenzaufbau sind zentrale Themen der Förderung. So unterstützen BMBF und Volkswagen-Stiftung beispielsweise Postdotorandinnen und Postdotoranden durch Fellowships und Nachwuchsgruppen. Darüber hinaus soll der Wissenstransfer durch geeignete langfristige Strukturen (dauerhafte Finanzierung, Langzeitarchivierung von Forschungsdaten, kompetentes Personal) getragen werden: Die Mehrheit der Programmausschreibungen formuliert den nachhaltigen Betrieb von Infrastrukturen als elementares Förderziel.

Tabellarische Übersicht der Förderaktivitäten

Die aufgeführten Förderaktivitäten geben einen Einblick in die Aktivitäten der genannten Mittelgeber und stellen keine abschließende Liste dar. Die Darstellung der EU-Förderaktivitäten beschränkt sich auf die Nennung der Förderlinie Forschungsinfrastrukturen und e-Infrastrukturen im Horizon2020-Programm, auf eine detaillierte Darstellung wurde wegen des Umfangs verzichtet.

Förderer	Förderprogramm/-richtlinie/-aktivität	Förderlinie/-rahmen [funding scheme]
BMBF	↗ <u>Vernetzen- Erschließen- Forschen. Allianz für universitäre Sammlungen</u> (2015): Erhöhung der Sichtbarkeit wissenschaftlicher Sammlungen an Universitäten durch Kooperationen und digitale Erschließung	Geistes-, Kultur- und Sozialwissenschaften
BMBF	↗ <u>Deutsches Netzwerk für Bioinformatik-Infrastruktur</u> (2015): Bereitstellung bioinformatischer Dienstleistungen und Trainings für die Lebenswissenschaften und die biomedizinische Grundlagenforschung	Gesundheitsforschung
BMBF	↗ <u>Förderkonzept Medizininformatik</u> (2015): Aufbau von Datenintegrationszentren an Universitätskliniken, Verbesserung von Forschung und Patientenversorgung durch IT-Lösungen	Gesundheitsforschung
BMBF	↗ <u>Verbund Forschungsdaten Bildung</u> (2013): Forschungsverbund der Forschungsdatenzentren GESIS, DIPF und IQB zur Sicherung des Zugangs vorhandener und künftiger Forschungsdaten	Bildungsforschung
BMBF	↗ <u>Die Sprache der Objekte</u> (2013): Erforschung musealer Objekte	Geistes-, Kultur- und Sozialwissenschaften
BMBF	↗ <u>eHumanities</u> (2013): Aufbau von Kompetenzzentren für die Digital Humanities	Geistes-, Kultur- und Sozialwissenschaften
BMBF	↗ <u>Deutsches Pflanzen Phänotypisierungsnetzwerk (DPPN)</u> (2012): Entwicklung im Hochdurchsatz arbeitender, automatischer Phänotypisierungsanlagen für Pflanzen	Bioökonomie
Bund-Länder-Förderung (GWK)	↗ <u>Nationale Kohorte</u> (2012): Nationale Langzeitstudie zur Erforschung von Volkskrankheiten mit Aufbau von zwei Forschungsdatenzentren	Bund-Länder-Vereinbarungen
Bund-Länder-Förderung (GWK)	↗ <u>Ausführungsvereinbarung über die gemeinsame Förderung von Forschungsbauten an Hochschulen einschließlich Großgeräten (AV-FuG)</u> (2007): Investitionen in die Entwicklung von Forschungsinfrastruktur, z. B. Hochleistungsrechner	Bund-Länder-Vereinbarungen
DFG	↗ <u>e-Research-Technologien</u> (2016) Auf- und Ausbau überregionaler, digitaler Informationsinfrastrukturen für einzelne Wissenschaftsbereiche oder als übergreifende Angebote; Schwerpunkte sind Technologien, Werkzeuge und Verfahren des Informationsmanagements	Wissenschaftliche Literaturversorgungs- und Informationssysteme (LIS)

DFG (u. a.)	Digging into Data Challenge (2016): Internationale Kooperationen (Beteiligung von mind. drei Ländern) im Bereich datenintensiver Forschung in den Geistes- und Sozialwissenschaften; Schwerpunkt: Nachnutzung bereits bestehender Infrastrukturen (Ausschreibung; erste Ausschreibungsrunde ohne Beteiligung der DFG)	Trans-Atlantic Platform (T-AP)
DFG	Forschungsdaten in der Praxis (2015-2016): Wissenschaftliche Nachnutzung bereits vorhandener Forschungsdatensätze aus etablierten Informationsinfrastruktureinrichtungen (einmalige Ausschreibung im Rahmen des Programms „Informationsinfrastrukturen für Forschungsdaten“)	LIS
DFG	Virtuelle Forschungsumgebungen (2015): Aufbau von virtuellen Arbeitsplattformen zur überregionalen Vernetzung von Forschenden (aufgegangen in: „e-Research-Technologien“)	LIS
DFG	Informationsinfrastrukturen für Forschungsdaten (2013): Entwicklung und Aufbau von Forschungsdatenrepositorien in Kooperation von Forschenden mit Infrastruktureinrichtungen	LIS
DFG	Erschließung und Digitalisierung (mind. seit 2007): Erschließung und Digitalisierung von Beständen aus wissenschaftlichen Bibliotheken und Archiven	LIS
DFG	Informationsmanagement und Informationsinfrastruktur in Sonderforschungsbereichen (INF) (2007): Aufbau von Datenbanken und Virtuellen Forschungsumgebungen in SFBs	koordinierte Programme: SFB/Transregio
EC	H2020 Excellent Science: Research infrastructure, including e-infrastructures (2014): In Anschluss an das Vorgängerprogramm FP 7 insbesondere Förderung der weiteren Integration von Informationsinfrastrukturen und der Nutzung von Synergiepotenzialen, Förderung von Open-Science-Projekten	1) Förderung von Verbundprojekten 2) Förderung von Koordinierungs- und Unterstützungsmaßnahmen
VW-Stiftung	Internationale Forschung in den Computational Social Sciences (2015): Sozialwissenschaftliche Forschung mit Daten, die durch die Nutzung neuer Medien entstanden sind; Vernetzung und Nachwuchsförderung	1) Fördermittel für internationale Workshops und Sommerschulen 2) Fördermittel für kooperative Forschungsvorhaben von PostdoktorandInnen
VW-Stiftung	Forschung in Museen (2008): Unterstützung musealer Forschungsaktivitäten durch Veranstaltungen und Projekte, Nachwuchsförderung im Bereich Sammlungsforschung	1) Fellowships für PostdoktorandInnen 2) Kooperative Forschungsprojekte an mittleren und kleinen Museen 3) Förderung von Workshops und Symposien

Stand: April 2016

[↗](#): Hinweis auf Hyperlink

E. Mitwirkende

E.1 RATSMITGLIEDER

Vertreter der wissenschaftlichen Nutzer

Prof. Dr. LARS BERNARD

Fakultät für Umweltwissenschaften, Technische Universität Dresden

Prof. Dr. Dr. h.c. FRIEDERIKE FLESS

Deutsches Archäologisches Institut und Freie Universität Berlin

Prof. Dr. FRANK OLIVER GLÖCKNER

Max-Planck-Institut für Marine Mikrobiologie und Jacobs University Bremen gGmbH

Prof. Dr. STEFAN LIEBIG

Fakultät für Soziologie, Universität Bielefeld

Prof. Dr. WOLFGANG MARQUARDT

Forschungszentrum Jülich

Prof. Dr. OTTO RIENHOFF

Institut für Medizinische Informatik, Georg-August-Universität Göttingen

Prof. Dr. JOACHIM WAMBSGANß

Zentrum für Astronomie der Universität Heidelberg (ZAH)

Prof. Dr. DORIS WEDLICH

Bereich Biologie, Chemie und Verfahrenstechnik, Karlsruher Institut für Technologie (KIT)

Vertreter der Einrichtungen

SABINE BRÜNGER-WEILANDT

FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur GmbH

Prof. Dr. THOMAS BÜRGER

Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden

Prof. Dr. PETRA GEHRING

Fachbereich Gesellschafts- und Geschichtswissenschaften, Technische Universität Darmstadt

Dr. GREGOR HAGEDORN

Museum für Naturkunde – Leibniz-Institut für Evolutions- und Biodiversitätsforschung

Prof. Dr. MICHAEL JÄCKEL

Universität Trier

Dr. MARGIT KSOLL-MARCON

Staatliche Archive Bayerns

Prof. Dr. KLAUS TOCHTERMANN

Deutsche Zentralbibliothek für Wirtschaftswissenschaften Kiel/Hamburg
und Christian-Albrechts-Universität zu Kiel

Prof. Dr. RAMIN YAHYAPOUR

Gesellschaft für Wissenschaftliche Datenverarbeitung mbH (GWDG)
und Georg-August-Universität Göttingen

Vertreter von Bund und Ländern

RÜDIGER EICHEL

Niedersächsisches Ministerium für Wissenschaft und Kultur

Dr. THOMAS GRÜNEWALD

Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen

Dr. STEFAN LUTHER

Bundesministerium für Bildung und Forschung

Dr. DIETRICH NELLE

Bundesministerium für Bildung und Forschung

Vertreter des öffentlichen Lebens

Dr. habil. REINHARD BREUER

Freier Journalist

Dr. h.c. ALBRECHT HAUFF

Thieme Verlagsgruppe

Dr. SIMONE REHM

TRUMPF GmbH und Co. KG (bis 12/2015)

ANDREA VOßHOFF

Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit

Vorsitz

Prof. Dr. OTTO RIENHOFF

Vorsitzender

SABINE BRÜNGER-WEILANDT

Stellvertretende Vorsitzende

E.2 GÄSTE UND EXTERNE EXPERTEN

Gäste

PETER BÜTTGEN

Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit

Dr. HANS-JOSEF LINKENS

Bundesministerium für Bildung und Forschung

Prof. Dr. NORBERT LOSSAU

Georg-August-Universität Göttingen

Dr. STEFAN WINKLER-NEES

Deutsche Forschungsgemeinschaft

Dr. h.c. PETER WITTENBURG

Max Planck Computing and Data Facility

Konsultation der AG Internationale Orientierung vom April 2016

Prof. Dr. JAN CARLSTEDT-DUKE

Karolinska Institutet, Schweden

Dr. RON DEKKER

Netherlands Organisation for Scientific Research, Niederlande

Prof. RNDr. LUDĚK MATYSKA

Masaryk University, Tschechische Republik

WALTER STEWART

Research Data Canada / Walter Stewart & Associates, Kanada

Dr. PAUL WONG

Australian National Data Service, Australien

E.3 GESCHÄFTSSTELLE

Dr. BARBARA EBERT

Leiterin

Dr. STEPHANIE HAGEMANN-WILHOLT

Dr. SVEN RANK

Dr. ILJA ZEITLIN

Referenten

SARAH BARUFFALDI

Assistenz

RFII – STRUKTUR UND ARBEITSGRUPPEN, STAND MAI 2016

