

Positionspapier „Research data at your fingertips“

der Arbeitsgruppe Forschungsdaten

[Februar 2015]

I. Vision 2025

„Research data at your fingertips“

Wissenschaftlerinnen und Wissenschaftler aller Disziplinen können auf alle Forschungsdaten einfach, schnell und ohne großen Aufwand zugreifen, um auf höchstem Niveau zu forschen und exzellente Ergebnisse zu erzielen. Sie können gemeinsam mit anderen arbeiten und ihre Forschungsergebnisse sicher aufbewahren. Forschungsdaten stehen 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 steigert die wissenschaftliche Reputation. Wissenschaftlerinnen und Wissenschaftler werden beim Sammeln, Erheben, Erfassen und beim Management ihrer Daten unterstützt.

Leicht nutzbare digitale Infrastrukturen sowie wissenschaftliche und technische Informationsspezialistinnen und -spezialisten unterstützen den vollständigen Forschungszyklus.

Die folgenden Ausführungen und Empfehlungen konzentrieren sich auf das Themenfeld „Forschungsdaten“. Die Arbeitsgruppe Forschungsdaten der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen stellt in diesem Arbeitspapier die aktuelle Situation im Themenfeld dar und benennt offene Herausforderungen, um damit eine Grundlage für notwendige Diskussionen und Überlegungen zum zukünftigen Umgang mit Forschungsdaten zu schaffen: <http://www.allianzinitiative.de/handlungsfelder/forschungsdaten.html>

Autorinnen u. Autoren: Michael Franke, Stefan Heinzl, Reiner Mauer, Dr. Janna Neumann, Dr. Heike Neuroth, Dr. Hans Pfeiffenberger, Henriette Senst, Andrea Sieberg, Dagmar Sitek, Claus Spiecker, Prof. Dr. Joachim Wambsganß, Dr. Danny Weber, Dr. Stefan Winkler-Nees

Lizenz: Alle Texte dieser Veröffentlichung sind unter einem Creative Commons Attribution 4.0 International (CC BY 4.0) Lizenzvertrag lizenziert. <http://creativecommons.org/licenses/by/4.0/>



Die Onlineversion dieser Publikation finden Sie unter:
<http://doi.org/10.2312/allianzfd.001>

II. Status Quo: Ein Zwischenstand

Das Themenfeld Forschungsdaten ist geprägt von einer großen Dynamik und Vielfalt. In einer Beschreibung des Status Quo müssen die folgenden wesentlichen Aspekte berücksichtigt werden:

1. Fachwissenschaftliche und strukturpolitische Betrachtung

Alle *Stakeholder* sind konfrontiert mit einem breiten und komplexen Aufgabenspektrum, das in weiten Teilen noch unklar und nicht spezifiziert ist. Konkrete Zuständigkeiten, wie etwa für die langfristige Archivierung von Forschungsdaten, sind häufig nicht bekannt oder noch ungeklärt bzw. nicht zugewiesen. Von besonderer Bedeutung sind dabei einerseits die Unterscheidung zwischen disziplinären und disziplinübergreifenden Aufgabenbereichen sowie die Frage, ob diese Aufgaben auf nationaler oder auch internationaler Ebene gebündelt werden sollten. Erste erfolgreiche Ansätze sind zum Beispiel disziplinübergreifende Dateninfrastrukturen¹ oder disziplinspezifische und internationale Forschungsdaten-Infrastrukturen².

Derzeitige Organisationsstrukturen sind in Deutschland stark geprägt von föderalen Gegebenheiten und der zum Teil heterogenen Landschaft der Wissenschaftsorganisationen, die teilweise aus einer (wissenschaftlichen) Konkurrenzsituation heraus agieren. Dies führt zu überlappenden oder auch nicht miteinander kompatiblen Entwicklungen.

Gleichwohl wurden in vielen Teilbereichen wichtige Fortschritte erreicht, wie etwa beim Aufbau von Angeboten zur Vergabe von persistenten Identifikatoren, die eine dauerhafte Referenzierbarkeit und Zitierbarkeit von Forschungsdaten ermöglichen oder beim Aufbau von universitären Beratungsangeboten zur Unterstützung beim Forschungsdatenmanagement. Diese sind wertvolle Bausteine bei der weiteren Entwicklung von Forschungsdaten-Infrastrukturen.

2. Bewusstsein für den professionellen Umgang mit Forschungsdaten

In manchen Disziplinen, wie z.B. den Geo- und Umweltwissenschaften oder den Sozialwissenschaften, hat sich bereits ein ausgeprägtes Bewusstsein für den Umgang mit Forschungsdaten etabliert und es wird unter Anwendung geeigneter Werkzeuge ein professionelles Datenmanagement betrieben. Meist hat sich eine fachspezifische Kultur des nachhaltigen Umgangs mit Daten im Sinne der Etablierung guter wissenschaftlicher Praxis entwickelt.

Zugleich wurden durch die Nachnutzung von Forschungsdaten wertvolle Ergebnisse im wissenschaftlichen Erkenntnisprozess erzielt. Die zunehmende Verfügbarkeit von Daten ermöglicht darüber hinaus in bestimmten Bereichen auch die Nachnutzung für gesellschaftlich relevante Fragestellungen außerhalb der Wissenschaft. So bilden z.B. Daten aus der Epidemiologie eine wichtige Grundlage für Erkenntnisse über die Verbreitung von Krankheiten sowie deren Gefahrenpotenzial und geeignete Präventionsmaßnahmen.

Eine zunehmende Bedeutung für die wissenschaftliche Forschung gewinnen Daten, die in nicht-wissenschaftlichen Kontexten entstehen. Hierzu gehören Daten, die auf gesetzlicher Grundlage oder im Rahmen von Verwaltungsprozessen entstehen (*Public Sector Information*), Daten aus sozialen Netzwerken, dem Gesundheitssektor oder *Citizen Science* Initiativen.

In etablierten, öffentlich finanzierten Informationseinrichtungen³ besteht ein unterschiedlich ausgeprägtes Bewusstsein im Hinblick auf den Umgang mit Forschungsdaten, aber auch eine be-

¹ European Data Infrastructure (EUDAT), <http://www.eudat.eu/> (02.02.2015)

² Z.B. EMBL-EBI, <http://www.ebi.ac.uk/> (02.02.2015) oder ELIXIR, <http://www.elixir-europe.org/> (02.02.2015); ESFRI-ERICs wie z.B. CEESDA, CLARIN, DARIAH, SHARE etc.

³ Archive, Bibliotheken, Museen, statistische Ämter etc.

ginnende Bereitschaft, sich damit auseinander zu setzen. Zahlreiche kommerzielle Unternehmen⁴ haben den wirtschaftlichen Wert von Daten ebenfalls erkannt.

3. Wissenschaftliche Organisationen

Viele wissenschaftliche Organisationen adressieren den Umgang mit Forschungsdaten, setzen ihn aber größtenteils noch in heterogener und unkoordinierter Art und Weise um. Die Allianz AG Forschungsdaten hat aus diesem Grund die „**Grundsätze zum Umgang mit Forschungsdaten**“ entwickelt, die die Vorstände aller Allianz-Organisationen 2010 verabschiedet haben⁵.

Allerdings ist vor allem die Finanzierung geeigneter Infrastrukturen für Forschungsdaten immer noch nicht geklärt. Die Förderorganisationen haben diese Anforderung erkannt und eine Reihe von Fördermaßnahmen entwickelt. Gemeinsam ist diesen Maßnahmen, dass sie nur im Rahmen befristeter Projekte wirken und damit den erforderlichen nachhaltigen Betrieb von Infrastrukturen nicht ermöglichen. Zugleich sind die verschiedenen Förderangebote nicht zwischen den Förderorganisationen abgestimmt.

Im Rahmen der Forschungsförderung muss seitens der Förderer anerkannt werden, dass Datenmanagement unabdingbarer Bestandteil von Forschungsprojekten ist und finanzieller Unterstützung bedarf.

Eine wichtige Voraussetzung für den Umgang mit Forschungsdaten sind 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.

Für die erfolgreiche Implementierung eines professionellen Umgangs mit Forschungsdaten braucht es auch politische Unterstützung. Hier sind bereits erste positive Schritte zu beobachten, wie die beginnende Bund/Länder Abstimmungen (GWK⁶) oder auch die Aktivitäten auf internationaler Ebene⁷ zeigen. Von einer Institutionalisierung von Forschungsdaten-Infrastrukturen sind wir jedoch noch weit entfernt. So sind z.B. rechtliche Rahmenbedingungen in weiten Bereichen noch nicht ausreichend geklärt.

Unentbehrlich für wissenschaftliches Datenmanagement sind Informationsspezialistinnen / Informationsspezialisten und -dienstleister. In der gegenwärtigen Praxis werden diese Aufgaben überwiegend durch Quereinsteigerinnen / Quereinsteiger übernommen; es fehlt der systematische Ausbau neuer Berufsfelder.

4. Agieren im internationalen Raum

Forschungsdaten sind nicht an nationale Grenzen gebunden. Zahlreiche Wissenschaftlerinnen und Wissenschaftler kooperieren international und erlangen gerade dadurch neue Erkenntnisse

⁴ Sog. Big Data Konzerne, forschende Unternehmen, Verlage etc., wie auch Unternehmensberatungen [McKinsey Global Institute: Big data: The next frontier for innovation, competition, and productivity, http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation (02.02.2015)]

⁵ Allianz Grundsätze zum Umgang mit Forschungsdaten, <http://www.allianzinitiative.de/handlungsfelder/forschungsdaten/grundsätze.html> (02.02.2015)

⁶ Vgl. Rat für Informationsinfrastrukturen, <http://www.gwk-bonn.de/index.php?id=205> (02.02.2015)

⁷ Z.B. G8+O5 Global Research Infrastructure, http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=international_level (02.02.2015); Research Data Alliance (RDA), <https://rd-alliance.org/> (02.02.2015)

und Reputation. Dies muss durch international anschlussfähige Infrastrukturen unterstützt werden. Erste positive Ansätze sind bereits erkennbar⁸, stehen jedoch vor zahlreichen Herausforderungen. So existieren nachhaltige Finanzierungsmodelle bisher nur in Ansätzen⁹.

III. Empfehlungen

Die Umsetzung der folgenden Empfehlungen trägt dazu bei, den Wissenschaftlerinnen und Wissenschaftlern in Deutschland in naher Zukunft eine adäquate Forschungsdaten-Infrastruktur bereitzustellen. Eine wesentliche Voraussetzung dafür ist die wissenschaftliche Anerkennung der Veröffentlichung von Forschungsdaten zur Nachnutzung (*citabile and accessible*) und deren Berücksichtigung in Evaluierungsverfahren von Personen und Institutionen.

1. Bundesweite disziplinspezifische Koordinationsgremien sind notwendig, die die jeweiligen Bedarfe feststellen, die notwendigen Infrastrukturkomponenten identifizieren, deren Auf- und Ausbau vorantreiben sowie relevante *Policies* und Standards setzen. Idealerweise sind in diesen Gremien Wissenschaft, Politik, Forschungsförderung und Infrastrukturen vertreten. Als ein Beispiel mit Vorbildcharakter kann der „Rat für Sozial- und Wirtschaftsdaten (RatSWD)“¹⁰ gesehen werden.
2. Vergleichbar mit Bibliotheken und ihrer Informationskompetenz bei der Literaturversorgung bedarf es für die nachhaltige Versorgung der Wissenschaft mit Forschungsdaten der Etablierung von vertrauenswürdigen (*trustworthy*) rechtsverbindlichen Organisationsstrukturen. Ideal wäre der Auf- und Ausbau von fachwissenschaftlichen Forschungsdatenzentren.
3. Dauerhafte und stabile Finanzierungsmodelle zur langfristigen Deckung der Entwicklungs- und Investitionskosten sowie der kontinuierlich anfallenden Betriebskosten sind zu entwickeln und umzusetzen. Eine wesentliche Komponente ist dabei eine angemessene Grundfinanzierung der Infrastruktur.
4. Forschung benötigt für die Nutzung und Publikation von Forschungsdaten klare und eindeutige rechtliche Rahmenbedingungen. Geregelt werden muss unter anderem die rechtssichere wissenschaftliche Nutzung von personenbezogenen Daten, urheberrechtlich geschützten und lizenzierten Materialien.
5. Um globale und vernetzte Forschung zu ermöglichen, sind Standardisierungen im technischen Bereich wie auch bei semantischen und syntaktischen Fragen notwendig. Internationale Koordination, Abstimmung und Implementierung stehen dabei im Vordergrund wie etwa die erfolgreichen Beispiele DFN/GEANT¹¹ (Netzinfrastruktur), INSPIRE¹² (geolokalisierte Daten) und DDI¹³ (Sozialwissenschaften) zeigen.
6. Der digitale Transformationsprozess erfordert eine zusätzliche Ausbildung der Fachwissenschaftlerinnen und Fachwissenschaftler. Neue Berufsfelder und Karrierewege sind durch die Einrichtung weiterer Studiengänge für Informationsspezialistinnen und Informationsspezialisten zu etablieren.



⁸ European Strategy Forum on Research Infrastructures (ESFRI), http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri (02.02.2015); Science Europe, <http://www.scienceeurope.org/> (02.02.2015)

⁹ Z.B. wurde für die ESFRI-Projekte eine eigene Rechtsform entwickelt: European Research Infrastructure Consortium (ERIC), http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric3 (02.02.2015)

¹⁰ Rat für Sozial- und Wirtschaftsdaten, <http://www.ratswd.de/> (02.02.2015)

¹¹ Deutsches Forschungsnetz, <https://www.dfn.de/> (02.02.2015) und Pan-European Data Network, <http://www.geant.net/> (02.02.2015)

¹² Infrastructure for Spatial Information in the European Community, <http://inspire.ec.europa.eu/> (02.02.2015)

¹³ Data Documentation Initiative, <http://www.ddialliance.org/> (02.02.2015)