

## Positionspapier

### **der Union der deutschen Akademien der Wissenschaften zur Schaffung einer Nationalen Forschungsdateninfrastruktur (NFDI)**

Der Rat für Informationsinfrastrukturen hat mit seinen Empfehlungen zum Forschungsdatenmanagement (Positionspapier „Leistung aus Vielfalt“, 2016) und seinen Überlegungen zum Einstieg in eine Nationale Forschungsdateninfrastruktur (Diskussionspapier „Wer bringt was mit?“, April 2017) wichtige Impulse zur Schaffung einer Nationalen Forschungsdateninfrastruktur gegeben. Die Wissenschaftsakademien halten die Einrichtung nachhaltiger Infrastrukturen für die Bereitstellung und Pflege von Forschungsdaten für dringend geboten. Dies gilt natürlich für alle Fachrichtungen, die digitale Daten von wissenschaftlich langfristiger Bedeutung erzeugen. Mit diesem Positionspapier nehmen die Akademien zum Vorstoß des RfII Stellung und fokussieren dabei speziell auf die Einrichtung einer NFDI für die Geisteswissenschaften, da dies für die Aktivitäten der Akademien von vorrangiger Bedeutung ist.

Die Wissenschaftsakademien in Deutschland sind sowohl Gelehrtenvereinigungen als auch Forschungseinrichtungen. Die Forschung der Akademien ist überwiegend geisteswissenschaftlicher Art. Dabei kommt den Akademien gerade dadurch, dass sie Langzeitforschungsprojekte durchführen, eine besondere Stellung innerhalb der geisteswissenschaftlichen Forschung zu. Mit ihren Editionen, Wörterbüchern, Lexika und Erschließungsprojekten leisten die Akademien wichtige Beiträge zur geisteswissenschaftlichen Grundlagenforschung. Schon vor Jahren haben die Akademien hierzu spezielle IT-Fachabteilungen aufgebaut bzw. betreiben sogar große wissenschaftliche Rechenzentren wie das Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften, die unter anderem für die Bereitstellung, Vernetzung, Pflege und Sicherung von Forschungsdaten und elektronischen Publikationen Sorge tragen. Indem die Akademien im eigenen Hause (und teilweise auch darüber hinaus) Infrastrukturdienste anbieten und damit de facto auch als Datenzentren fungieren, sind sie sowohl mit wesentlichen Anforderungen aus der Nachfrageperspektive als auch mit den Herausforderungen aus der Angebotsperspektive vertraut.

Wie in den Beiträgen des RfII formuliert, vertreten auch die Akademien die Auffassung, dass der Aufbau einer nationalen Forschungsdateninfrastruktur ebenso wie deren späterer, auf Dauer und Nachhaltigkeit ausgerichteter Betrieb ein nachfragegetriebener Prozess sein muss. Im Folgenden wird daher zunächst auf die Nachfrageperspektive eingegangen. Zugleich sehen sich die Akademien in der Pflicht, sich sowohl an der Diskussion als auch dem Aufbau eines entsprechenden Angebotes zu beteiligen.

## **I. Anforderungen an eine Forschungsdateninfrastruktur für die Geisteswissenschaften (Nachfrageperspektive)**

### **1. Dauerhaftigkeit und grundlegende Anforderungen**

Aus Sicht der Akademien besteht für die Geisteswissenschaften ein dringender Bedarf an einer Infrastruktur, die jedem geisteswissenschaftlichen Forschungsprojekt die langfristige Sicherung seiner digital vorliegenden Daten und Ergebnisse ermöglicht. Sicherung ist hier im Sinne einer *Langzeitarchivierung* zu verstehen, die insbesondere auch die Möglichkeit der Archivierung auf unbegrenzte Zeit vorsieht. Letzteres ist speziell für solche Forschungsdaten notwendig, die nicht ein zweites Mal erhoben werden können (z.B. für Daten aus archäologischen Ausgrabungen) und die in der weiteren Forschung kontinuierlich genutzt werden sollen (z.B. digitale Repräsentationen von Objekten des kulturellen Erbes).

Daneben muss eine Forschungsdateninfrastruktur (FDI) auch die nachhaltige Bereitstellung digitaler Forschungsdaten und -ergebnisse nach Abschluss von Forschungsprojekten durch den Betrieb entsprechender Softwaresysteme (Online-Repositorien, digitale Publikationen, webbasierte Fachinformationssysteme und Portale) nach den FAIR-Prinzipien (Findable, Accessible, Interoperable, Re-Usable) ermöglichen. Die FAIR-Prinzipien beinhalten die offene Lizenzierung der Forschungsdaten und offene Schnittstellen für den Zugriff darauf. Anders als die Archivierung, muss die *Langzeitbereitstellung* nicht unbedingt zeitlich unbegrenzt funktionieren, sondern kann je nach Bedarf der Fachgemeinschaft im Verhältnis zu den Bereitstellungskosten zeitlich beschränkt und ggf. auf eine Archivierung der Daten zurückgefahren werden.

In jedem Fall aber sollte die Datenbereitstellung so gestaltet werden, dass sie die digitale Erschließung geisteswissenschaftlicher Daten durch die zukünftige Forschung unterstützt. Dies erfordert eine sorgfältige Datenkuratierung, nachnutzbare Annotierungsstandards und ggf. auch die Bereitstellung und Pflege von Softwarewerkzeugen.

### **2. Lebendige Weiterentwicklung und Vielfalt**

Die Forschungsdateninfrastruktur muss mit der hohen Entwicklungsdynamik sowohl der Technik, die sie einsetzt, als auch der wissenschaftlichen Disziplinen, die sie bedient, Schritt halten können. Dazu muss sie auf den kontinuierlichen Ausbau und die Integration zukünftiger Bedarfe und Technologien vorbereitet sein.

Eine Bedingung *sine qua non* ist, dass die Forschungsdateninfrastruktur auch fachbereichsspezifischen Anforderungen gerecht wird. Eine FDI, die das nicht leistet, würde ihren Zweck verfehlen. Vor diesem Hintergrund scheint eine Grobeinteilung in naturwissenschaftliche und geisteswissenschaftliche NFDI-Konsortien sinnvoll. Auf der anderen Seite würde der Aufbau einer eigenen FDI für jedes Fach unnötig viele Ressourcen binden und könnte für den Datenaustausch zwischen Fachdisziplinen sogar kontraproduktiv sein. Auch grobeingeteilte Konsortien müssten ihre Lösungen soweit aufeinander abstimmen, dass die interdisziplinäre Zusammenarbeit dort, wo sie sinnvoll ist, wie z.B. in der historischen Geographie, nicht an technische Barrieren stößt.

Wir gehen von der Prämisse aus, dass zwischen den geisteswissenschaftlichen Fächern hinsichtlich der Anforderungen an das Datenmanagement hinreichend viele Übereinstimmungen bestehen, so dass ein Konsortium so gut wie alle Fächer sinnvoll bedienen kann. Die fachspezifischen Bedürfnisse können durch eine subsidiäre Organisationsform eingefangen werden, etwa durch ein Netzwerk von Datenrepositorien, die eigene fachliche Schwerpunktsetzungen vornehmen, aber gemeinsamen fächerübergreifenden Standards genügen. Nicht weniger wichtig als die fächerübergreifende Anschlussfähigkeit ist die überregionale Anschlussfähigkeit an europäische und internationale Infrastrukturen.

Der Berücksichtigung von Nachfragebedürfnissen kann am besten Rechnung getragen werden, wenn es für die Nachfragenden eine Wahlmöglichkeit zwischen alternativen Angeboten – und damit Wettbewerb – gibt. Naturgemäß ist das bei Infrastrukturangeboten nicht immer zu realisieren. Dort, wo das möglich erscheint (z.B. bei dem Angebot von Softwarewerkzeugen) sollte Wettbewerb realisiert werden. Wo nicht, sollten die Nachfragenden über Mitsprache- und Mitgestaltungsmöglichkeiten verfügen, um der Gefahr von Verkrustung und Monopolisierung entgegenzuwirken.

### **3. Vermittlung von Wissen und Befähigungen**

Besonders in den Geisteswissenschaften ist derzeit eine zunehmende Kluft zu beobachten zwischen denen, die digitale Methoden kennen und diese nutzen, und denen, die keinen Zugang dazu haben, oft, weil nicht einmal bekannt ist, welche technischen Möglichkeiten es überhaupt gibt.

FDI-Einrichtungen sollten daher auch die Verbreitung von Know-How und den Erwerb von Fähigkeiten zum richtigen Umgang mit digitalen Forschungsdaten in den Geisteswissenschaften aktiv fördern, u.a. durch eine geeignete Nachwuchsförderung, aber auch durch die Unterstützung von Fachbereichen bei der Integration entsprechender Lehrbausteine in ihre Curricula. Idealerweise sollten Studierende schon in den Einführungen zum wissenschaftlichen Arbeiten mit dem Umgang mit Forschungsdaten und Forschungsdateninfrastrukturen vertraut gemacht werden – ganz so, wie sie bisher schon die Nutzung von Bibliotheken und Bibliographien in den Einführungsveranstaltungen lernen.

## **II. Die Aktivposten der Akademien für eine geisteswissenschaftliche Forschungsdateninfrastruktur (Angebotsperspektive)**

Forschungsdateninfrastrukturen bedürfen eines stabilen institutionellen Rahmens. Durch ihre bis über 300 Jahre zurückreichende Geschichte bieten die Akademien einen solchen stabilen institutionellen Rahmen. Als unmittelbare Einrichtungen der Bundesländer sind sie fest in das föderale Wissenschaftssystem Deutschlands integriert. Gleichzeitig sind sie mit der Akademienunion auch auf Bundesebene organisiert. Gerade im digitalen Bereich findet in diesem Rahmen eine intensive und regelmäßige Zusammenarbeit zwischen den Akademien statt. Über den All European Academies Verband (ALLEA) sind die deutschen Akademien zudem europaweit vernetzt. Die Akademien sehen sich daher sowohl in der Lage als auch in der Verantwortung, einen Beitrag zum Aufbau einer Nationalen Forschungsdateninfrastruktur für die Geisteswissenschaften zu leisten.

Da die Akademien überwiegend Langzeitforschungsvorhaben mit Laufzeiten von bis zu 25 Jahren betreiben, sind sie mit den Anforderungen einer langfristig angelegten Erstellung und Pflege von geisteswissenschaftlichen Forschungsdaten wohlvertraut. Insbesondere mit ihren Wörterbüchern, Erschließungsprojekten (in Form digitaler Kataloge und Nachschlagewerke) und digitalen Editionen bringen die Akademien Grundlagenwerke hervor, die die Forschungsinfrastruktur der entsprechenden geisteswissenschaftlichen Fächer nachhaltig stärken.

Viele Akademien verfügen bereits seit Jahren über Fachabteilungen, die die oben genannten Bedarfe im eigenen Haus oder über etablierte Kooperationen erfüllen. Dabei haben sich die Akademien umfangreiche Kenntnisse und Fähigkeiten für die digitale Erschließung, Bereitstellung und Pflege geisteswissenschaftlicher Forschungsdaten erarbeitet und passgenaue Softwarewerkzeuge und Dienste dafür bereitgestellt. Schon heute betreiben die Akademien aktives Forschungsdatenmanagement auch für Vorhaben, deren Förderung beendet ist. Sie verfügen bereits über Übernahme- und Verstetigungsroutinen und sind de-facto-Datenzentren, die Daten auf Dauer kuratieren und anbieten. Eine derzeit noch nicht vorhandene personelle und finanzielle Ausstattung vorausgesetzt, könnten sie diese Dienste auch anderen dauerhaft anbieten und ihr vorhandenes Know-How sowie entsprechende Werkzeuge im Rahmen von Beratungsprojekten an andere Institutionen weitergeben. Darüber hinaus sind die Akademien in der Lage, Dienste der von ihnen betriebenen Rechenzentren und anderen Serviceanbietern für die Langzeitarchivierung und -bereitstellung von Forschungsdaten mit einzubinden.

Durch ihre zahlreichen geisteswissenschaftlichen Forschungsprojekte sind die Akademien gerade auch mit den fachspezifischen Bedürfnissen und Anforderungen an ein Forschungsdatenmanagement gut vertraut. Sie können darauf ebenso eingehen, wie sie durch die fachliche Breite ihrer Forschungsaktivitäten fächerübergreifende Synergieeffekte beim Forschungsdatenmanagement freisetzen können.

In der zusammenführenden Betreuung von dauerhaft vorzuhaltenden Forschungsdaten aus verschiedenen zentralen Themenfeldern der Geisteswissenschaften können die Akademien Materialcluster bilden, die inhaltsreich, dauerhaft und stabil sein werden und zu einem Nutzungsmehrwert sonst verstreuter Daten führen. Die Akademien erforschen aktiv die Möglichkeiten der Vernetzung von Forschungsdaten (z.B. im Rahmen des GeRDI-Projekts) und tragen durch die Entwicklung und den Einsatz interoperabler Standards in Akademienvorhaben (z.B. TEI, DTA-Basisformat, RDF/LOD) sowie durch die Bereitstellung von Vernetzungsdiensten (z.B. correspSearch, IncipitSearch) zu einer einfacheren Nachnutzung und einer nachhaltigen Erschließung geisteswissenschaftlicher Forschungsdaten bei.

Die Akademien sind im Bereich der digitalen Grundlagenforschung (Digitale Editorik, digitale Lexikographie, digitale Musikwissenschaften u.v.w.) eng mit Universitäten, Hochschulen und außeruniversitären Forschungseinrichtungen verbunden. Ein Teil der Forschungsvorhaben der Akademien wird an oder in Kooperation mit Universitäten durchgeführt. Die Akademien unterstützen den digitalen Kompetenzaufbau und die praxisorientierte digitale Ausbildung des wissenschaftlichen Nachwuchses auf vielfältige Weise. Sie veranstalten Tagungen, Symposien und Workshops, bieten Lehrveranstaltungen an und beteiligen sich an Studiengängen im Bereich der digitalen Methodik.

### **III. Ausblick**

Der Aufbau einer verlässlichen Forschungsdateninfrastruktur für die Geisteswissenschaften bildet den Schlüssel für den Einsatz digitaler Methoden in den Geisteswissenschaften. Ohne eine leistungsfähige Forschungsdateninfrastruktur wird der Einsatz von digitalen Verfahren in den Geisteswissenschaften immer mit der Sorge um den Verlust oder das Unbrauchbarwerden des digital gespeicherten Wissens belastet sein. Die Chancen, die digitale Methoden der geisteswissenschaftlichen Forschung eröffnen, lassen sich nur mit einer auf Langfristigkeit ausgelegten Forschungsdateninfrastruktur voll ausschöpfen. Die Wissenschaftsakademien in Deutschland plädieren nachdrücklich für die Einrichtung einer Forschungsdateninfrastruktur, die dem Bedarf in den Geisteswissenschaften genügt und deren spezifischen Anforderungen Rechnung trägt. Die Akademien sind bereit, einen Beitrag zu deren Aufbau und dauerhaftem Betrieb zu leisten. Sie erwarten, dass eine solche Forschungsdateninfrastruktur das digitale kulturelle Erbe nachhaltig sichern, zum Türöffner für die Weiterentwicklung und breite Nutzung digitaler Methoden und Technologien in den Geisteswissenschaften werden und die geisteswissenschaftliche Spitzenforschung in Deutschland und darüber hinaus signifikant befruchten wird.

#### **Kontakt:**

Prof. Dr. Martin Grötschel

Sprecher der NFDI-Arbeitsgruppe der Union der deutschen Akademien der Wissenschaften  
und Präsident der Berlin-Brandenburgischen Akademie der Wissenschaften

Tel.: 030/20 37 06 45/-20

E-Mail: p@bbaw.de

#### **Pressekontakt:**

Dr. Annette Schaeffgen

Leiterin des Berliner Büro der Union der deutschen Akademien der Wissenschaften,  
Presse- und Öffentlichkeitsarbeit

Tel.: 030 / 325 98 73-70

E-Mail: schaeffgen@akademienunion-berlin.de