

Forschungsdatenstrategie der Hochschulen Baden-Württembergs

September 2023

Erarbeitet von

Prof. Dr. Vincent Heuveline, Universität Heidelberg

Oliver Kohl-Frey, Universität Konstanz

Prof. Dr. Thomas Walter, Universität Tübingen

im Auftrag der Leiterinnen und Leiter der Rechenzentren der Universitäten des Landes Baden-Württemberg (ALWR-BW)

Dr. Marianne Dörr, Universität Tübingen

Dr. Sabine Gehrlein, Universität Mannheim

Dr. Arne Upmeyer, Karlsruher Institut für Technologie

im Auftrag der Direktorinnen und Direktoren der Universitäts- und Landesbibliotheken Baden-Württemberg (AGBibDir-BW)

Dr. Renke Siems, Ministerium für Wissenschaft, Forschung und Kunst (ständiger Gast)

unter Beteiligung von

Valerie Boda, Karlsruher Institut für Technologie

Livia Gertis, Universität Konstanz

Dr. Sophie G. Habinger, Universität Konstanz

Dr. Cora F. Krömer, Karlsruher Institut für Technologie

in Abstimmung mit den Hochschulen für angewandte Wissenschaften des Landes Baden-Württemberg

mit Zustimmung der Landesrektorenkonferenz (LRK) der Universitäten Baden-Württemberg

Revision 1.0, 29.09.2023

Inhaltsverzeichnis

Präambel	4
Executive Summary	5
Forschungsdaten(management) in Baden-Württemberg	6
Was ist Forschungsdatenmanagement (FDM)?	6
Aktueller Stand des Forschungsdatenmanagements in BW	8
Herausforderungen und Perspektiven	9
Handlungsfeld Personen	11
Ausgangslage und Perspektive	11
Organisation	13
Empfehlungen	15
Handlungsfeld Governance	16
Ausgangslage und Perspektive	16
Empfehlungen	21
Handlungsfeld Technik	22
Ausgangslage und Perspektive	23
Themenfelder	25
Empfehlungen	28

Präambel

Der digitale Wandel stellt das Land Baden-Württemberg vor langfristige Herausforderungen: Immer mehr Daten müssen angemessen aufbereitet, sicher verarbeitet, verwaltet und gespeichert werden, um einen bestmöglichen Nutzen für Wissenschaft, Wirtschaft und Gesellschaft zu erbringen. Hochschulen und anderen Forschungseinrichtungen kommt in diesem Zusammenhang eine besondere Rolle zu; sie sind Pioniere im Aufbau von Datenökosystemen. Ihre Forschungsdaten bilden die Grundlage für wissenschaftlichen Fortschritt und Innovation. Gutes Forschungsdatenmanagement unterstützt dabei nicht nur Forschungsprozesse im Umfeld öffentlicher Einrichtungen. Die im Umgang mit Daten entstehenden Erkenntnisse können auch gewinnbringend in Wirtschaft und Verwaltung genutzt werden. Wissenschaftliches Forschungsdatenmanagement ist daher einer der Grundpfeiler für eine gelingende Digitalisierung und muss systematisch unterstützt und ausgebaut werden.

Baden-Württemberg nimmt sich der Herausforderungen durch die Digitalisierung seit vielen Jahren an: Mit dem Fachkonzept *E-Science – Wissenschaft unter neuen Rahmenbedingungen* sowie der Landesdigitalisierungsstrategie *digital@bw* leitete das Land den digitalen Wandel bewusst ein und gestaltet diesen mit *digital@LÄND* aktiv weiter. Die vorliegende Forschungsdatenstrategie schließt unmittelbar an die genannten Konzepte an. Sie wurde im Auftrag des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK) von Leitungen aus Universitätsbibliotheken und -rechenzentren des Landes entwickelt und im Kreis der Prorektorinnen und Prorektoren Forschung, in der Landesrektorenkonferenz sowie im Steuerkreis Digitale Forschungsinfrastruktur BW beraten. Bezogen auf verschiedene Aspekte des Forschungs(daten)prozesses ergänzen sich die Forschungsdatenstrategie und das Rahmenkonzept zum High Performance Computing (HPC / DIC) nahtlos.

Die Forschungsdatenstrategie ist ein wesentlicher Baustein für die Weiterentwicklung der Digitalisierungs- und Datenstrategie Baden-Württembergs. Denn eine wegweisende und ressourcenschonende Strategie zu Forschungsdaten stärkt auch alle weiteren, digital getriebenen Entwicklungen im Land, wie etwa die im Aufbau befindliche Spitzenforschung auf dem Gebiet der Künstlichen Intelligenz (KI) und die zukunftsorientierten Entwicklungen im Cyber Valley. Gleichzeitig gewinnt das Forschungsdatenmanagement durch die jeweils neuesten Entwicklungen auf dem Gebiet der KI-Forschung weiter an Qualität.

Ziel der Forschungsdatenstrategie der Hochschulen ist es, die Chancen der datenbasierten Transformation in der Forschung zu nutzen. Die Strategie empfiehlt – auf Basis einer Analyse des Ist-Zustands der Forschungsdateninfrastruktur in Baden-Württemberg – den wachsenden Anforderungen an das Forschungsdatenmanagement mit Lösungen in den Handlungsfeldern Personal, Governance und Technik zu begegnen. Hierfür werden Handlungsoptionen vorgeschlagen, die eine nachhaltige Entwicklung des Forschungsdatenmanagements in Baden-Württemberg fördern; d. h. auf langfristigen Bestand und einen ökonomischen Umgang mit personellen, finanziellen und ökologischen Ressourcen ausgerichtet sind.

Mit der vorliegenden Strategie wird eine gemeinsame Linie zum Forschungsdatenmanagement für alle Einrichtungen im Bereich des MWK entwickelt. Auf den Handlungsempfehlungen aufbauend wird die Forschungsdatenstrategie durch ein Umset-

zungskonzept ergänzt werden, welches konkrete Maßnahmen zur Zielerreichung benennt. Insgesamt positioniert sich das Land Baden-Württemberg damit im nationalen wie internationalen Kontext weiterhin als führender Standort innovativer, datengetriebener Forschung.

Die Forschungsdatenstrategie wurde im Auftrag des Ministeriums für Wissenschaft, Forschung und Kunst von einer Schreibgruppe aus dem Arbeitskreis der Leiter der wissenschaftlichen Rechenzentren (ALWR) und der Arbeitsgemeinschaft der Bibliotheksdirektorinnen und -direktoren der wissenschaftlichen Bibliotheken (AG BibDir) geschrieben. Die Strategie wurde im Sommer 2023 in den jeweiligen Arbeitskreisen der CIO der Universitäten, der Prorektorinnen und Prorektoren, Vizepräsidentinnen und Vizepräsidenten Forschung sowie in der Landesrektorenkonferenz diskutiert und dort jeweils zustimmend zur Kenntnis genommen.

Executive Summary

Die Forschungsdatenstrategie der Hochschulen Baden-Württembergs wird anhand von drei Handlungsfeldern dargestellt: Personen, Governance und Technik. In allen drei Feldern existiert Handlungsbedarf, will das Land seine derzeitige starke Stellung bei Forschungsdaten behalten.

Personen

Forschungsdatenmanagement ist ein dynamischer Bereich, der wissenschaftlich befördert, aber auch wissenschaftsunterstützend begleitet werden muss – beides erfordert qualifiziertes Personal. Durch die Zunahme datenorientierter Wissenschaft entsteht ein stetig wachsender Bedarf an personellen Ressourcen und Kompetenzen. Dieser kann mit bereits jetzt bestehenden Kapazitäten nicht nachhaltig gedeckt werden. Darüber hinaus wird die Langfristaufgabe Forschungsdatenmanagement derzeit noch maßgeblich durch temporäre Projektstrukturen getragen, innerhalb derer Personal gegebenenfalls nicht ausreichend qualifiziert und eingearbeitetes Personal nicht gehalten werden kann. Diesem Defizit an allgemeiner und fachspezifischer Expertise kann durch die Einrichtung langfristig finanzierter, lokaler Stellen begegnet werden. Grundlage ist die Etablierung von konkreten Berufsbildern, Qualifikationswegen und gegebenenfalls Studiengängen. Diese sollten im Land Baden-Württemberg vorangetrieben und koordiniert werden, um nach innen den Aufbau von effizienten, kooperativen Strukturen zu ermöglichen und nach außen die Anschlussfähigkeit an andere Bundesländer und die Nationale Forschungsdateninfrastruktur (NFDI) zu gewährleisten.

Governance

Mit einer erfolgreichen Governance werden steuernde und regelnde Strukturen der Forschungsdatenlandschaft Baden-Württemberg zusammengeführt und dadurch die Datensouveränität des Landes strategisch gestärkt. Dies gilt für die Entwicklung kooperativer Lösungen und dem damit verbundenen effizienteren Einsatz von Ressourcen ebenso wie für

die Entwicklung von verbindlichen Regeln im rechtssicheren Umgang mit Forschungsdaten. Hier weist das Land jedoch Leerstellen auf. Um allen Hochschulen Zugang zu einer adäquaten Forschungsdateninfrastruktur in Form von personellen Ressourcen und IT-Lösungen zu bieten, müssen die vorhandenen Strukturen nachhaltig vernetzt, ergänzt und strategisch gefördert werden. Deshalb sollten einrichtungsübergreifende Strukturen entstehen, die lokale Stellen, wo angemessen, ergänzen und entlasten. Neben einer derzeit nur als Projekt finanzierten zentralen Koordinierungs- und Kontaktstelle für alle FDM-Akteure und ihre Aktivitäten, sollten auch zentrale Stellen für spezialisiertes Wissen (juristische FDM-Kompetenz, Datentreuhand, Datenethik) eingerichtet werden. Zur Beratung auf politischer Entscheidungsebene wird die Einsetzung eines Landesdatenrates vorgeschlagen.

Technik

Zuverlässige Infrastruktur ist gleichermaßen Voraussetzung wie auch Ergebnis guter Governance und kompetenter Personalressourcen. Mit der BaWü-Datenföderation kann Baden-Württemberg erfolgreiche IT-Strukturen vorweisen. Um aber dem steigenden Bedarf und den wachsenden Anforderungen seitens der Wissenschaft nachkommen zu können, müssen diese mit Blick auf künftige Herausforderungen ausgebaut werden. Dies umfasst sowohl technische Lösungen (z. B. die Ermöglichung verlustfreien Datentransfers) als auch organisatorische (z. B. Aufbau von Datentreuhandstellen). Die Verarbeitung, Speicherung und Langzeitarchivierung von Forschungsdaten wird dauerhafte Investitionen erfordern, um die Datensouveränität sicherzustellen. Dies meint neben der Bereitstellung und Instandhaltung von technischer IT-Infrastruktur auch die Entwicklung innovativer Software-Lösungen für den Umgang mit Forschungsdaten. Übergreifende Strukturen, wie die NFDI, müssen durch regionale Angebote der Hochschulen ergänzt werden, um den tatsächlichen Bedarf und auch weitere Fachdisziplinen abzudecken und damit gleichzeitig die Forschung in Baden-Württemberg zu fördern. Lösungen hierfür wären zum Beispiel der Aufbau von regionalen Forschungsdatenrepositorien oder auch die Einrichtung eines verteilten Landesdatenarchivierungszentrums. Essentiell ist der Ausbau kooperativer und regionaler Lösungsansätze auch, um die Anschlussfähigkeit kleinerer Einrichtungen an landeseigene Infrastrukturen sicherzustellen.

Forschungsdaten(management) in Baden-Württemberg

Was ist Forschungsdatenmanagement (FDM)?

In der Wissenschaft wird gemessen, gesammelt und ausgewertet. Die im Zuge wissenschaftlichen Arbeitens entstehenden Forschungsdaten liegen mittlerweile meist (jedoch nicht immer) in digitaler Form vor und sind sehr heterogen: Sie stammen beispielsweise aus Beobachtungen, Experimenten, Simulationsrechnungen, Erhebungen oder Befragungen. Zu Forschungsdaten zählen darüber hinaus auch textuelle und andere Quellen, (amtliche) Statistiken und deren Auswertungen. Dementsprechend vielfältig sind auch die Datenformate: Das Spektrum reicht von entschlüsselten Pflanzengenomen über Daten in (elektronischen)

Laborbüchern bis zu Social Media Posts; von der einfachen Textdatei über Bild- und Audioformate jeglicher Größe bis zur Forschungssoftware, mit der die Ergebnisberechnung stattgefunden hat.

Neben diesen Primärdaten braucht es sogenannte Metadaten, die Kontextinformationen zu Daten und Datensätzen liefern. Diesen kommt eine zentrale Rolle für das Auffinden, Verstehen und Nachnutzen von Forschungsdaten zu: Sie liefern grundlegende Informationen (z. B. zu Urhebern etc.), beschreiben aber auch den Forschungsprozess genauer, in dem die Daten erhoben bzw. verarbeitet wurden. Erst durch die Verknüpfung mit Metadaten werden Primärdaten für die weitere Forschung wertvoll – bei der Reproduktion von Ergebnissen ebenso wie bei der Suche nach und der Nutzung von bereits vorhandenen Datensätzen. Welche fachwissenschaftlichen Metadaten für die Nachnutzbarkeit von Forschungsdaten notwendig sind, unterscheidet sich in den verschiedenen wissenschaftlichen Fachbereichen teilweise sehr stark.

Forschungsdaten können in fachspezifischen oder generischen, d. h. vom wissenschaftlichen Fachgebiet unabhängigen, Repositorien archiviert und zugänglich gemacht werden. Es gilt dabei auch für Datensätze immer häufiger die Vorgabe von Fördergebern und Hochschulpolitik, dass Daten möglichst offen und FAIR¹ zur Verfügung stehen sollten, sofern keine datenschutz- und urheberrechtlichen Gründe dagegen sprechen. Eine besondere Herausforderung ist in diesem Zusammenhang die Langzeitarchivierung von Forschungsdaten und Forschungssoftware: Sollen Daten über einen Zeitraum von mehr als zehn Jahren gelagert werden, ist darauf zu achten, dass sie auch wieder ausgelesen und verwendet werden können. Neben Daten und Software-Code müssen in diesen Fällen daher auch weitere Informationen erhalten bleiben, um auch nach Jahren noch eine neuerliche Verwendung und Interpretation der Forschungsdaten zu ermöglichen.

Der Umgang mit Forschungsdaten, die Dokumentation von Daten und deren Verfügbarmachung zur Nachnutzung wird unter dem Begriff Forschungsdatenmanagement zusammengefasst. Gutes Forschungsdatenmanagement erleichtert den Forschungsalltag und ist essentieller Bestandteil der guten wissenschaftlichen Praxis²: Die nachvollziehbare Aufbereitung von Forschungsdaten ermöglicht die bessere Überprüfung von Forschungsergebnissen im Rahmen von Reviewprozessen und Replikationsstudien. Forschungsdatenmanagement ist daher auch integraler Bestandteil von Open Science, die sich für möglichst barrierefreie Nachnutzungsmöglichkeiten von Forschungsergebnissen einsetzt. In standardisierter Form abgelegte Forschungsdaten können für zukünftige Fragestellungen genutzt werden; doppelte Arbeit wird vermieden. Ziel des Forschungsdatenmanagements ist es, Daten FAIR zu machen, und damit ihr volles Potenzial für aktuelle und zukünftige Forschung auszuschöpfen bzw. zu erhalten.

¹**FAIR**-Prinzipien: Forschungsdaten sollten auffindbar (**F**indable), zugänglich (**A**ccessible), interoperabel (**I**nteroperable) und wiederverwertbar (**R**eusable) gespeichert werden.

²Vgl. den Kodex der Deutschen Forschungsgemeinschaft (DFG) „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“, <https://doi.org/10.5281/zenodo.3923601>.

Aktueller Stand des Forschungsdatenmanagements in BW

Die baden-württembergische Forschungspolitik hat bereits früh damit begonnen, das Forschungsdatenmanagement systematisch zu fördern. Wichtige Bausteine hierfür waren die Konzepte *bwDATA*³ und speziell das 2014 veröffentlichte Fachkonzept *E-Science – Wissenschaft unter neuen Rahmenbedingungen*, das als eines der ersten Konzepte dieser Art die Verfügbarmachung von digitaler Information für die Forschung als infrastrukturelle Grundversorgungsaufgabe der Hochschulen in Baden-Württemberg benennt: Forschungsdatenmanagement wird darin als eines von fünf zentralen Handlungsfeldern identifiziert, in denen Infrastruktur und Governance-Strukturen aufgebaut werden müssen. Im Rahmen der genannten Konzepte – ebenso wie im Zusammenhang mit der Landesdigitalisierungsstrategie *digital@bw* – förderte das Land eine ganze Reihe von Projekten zum Forschungsdatenmanagement sowie zu virtuellen Forschungsumgebungen⁴. Besonders hervorzuheben sind in diesem Zusammenhang die vier Science Data Center (SDC) die ab 2019 für unterschiedliche Fachdisziplinen Werkzeuge und Standards zur Aufbereitung und Speicherung von Forschungsdaten entwickelten.⁵ Durch die in den SDC gewonnene Expertise leistet Baden-Württemberg einen wesentlichen Beitrag zum Aufbau der Nationalen Forschungsdateninfrastruktur (NFDI).

Das Land Baden-Württemberg fördert außerdem die Vernetzung zum Thema Forschungsdatenmanagement – auch national und international: Das Projekt *Landesinitiative bwFDM*⁶ erarbeitet neue Formate für Wissenstransfer und Kooperationen zwischen baden-württembergischen FDM-Expertinnen und -Experten. Zugleich wird eine Vernetzungsstruktur für den Austausch mit FDM-Akteuren auf nationaler Ebene (z. B. andere FDM-Bundeslandinitiativen, NFDI) organisiert. Die Informationsplattform *forschungsdaten.info* bietet seit 2016 eine stetig wachsende, aktuelle Informationszusammenstellung zu den Themen Forschungsdatenmanagement und Forschungsdateninfrastruktur; das Redaktionsteam bündelt FDM-Expertise aus dem gesamten D-A-CH-Raum. Und nicht zuletzt bietet die Konferenz *E-Science-Tage* der Auseinandersetzung mit dem Thema Forschungsdatenmanagement eine Plattform.

Die standortübergreifende Auseinandersetzung mit Herausforderungen des Forschungsdatenmanagements an baden-württembergischen Universitäten wurde von diesen durch die Gründung des Arbeitskreises Forschungsdatenmanagement (AK FDM) im Jahr 2018 adressiert. Empfehlungen und Handreichungen des Gremiums entfalten durch ihr Einbringen in die Gremien der Bibliotheksdirektorinnen und -direktoren (AG BibDir) sowie der Rechenzentrumsleitungen in Baden-Württemberg (ALWR) auch auf hochschulpolitischer Ebene Wirkung.

³Phase I: 2013-2014 (<http://dx.doi.org/10.15496/publikation-21188>), Phase II: 2015-2019 (<http://dx.doi.org/10.15496/publikation-21187>), Phase III: 2020-2024 (<http://dx.doi.org/10.15496/publikation-55923>).

⁴Virtuelle Forschungsumgebungen sind Softwarelösungen bzw. Plattformen, die die ortsunabhängige Zusammenarbeit zwischen Wissenschaftlerinnen und Wissenschaftlern ermöglichen.

⁵Wirtschaftswissenschaften (BERD@bw), Lebenswissenschaften (BioDaten), Chemie, makromolekulare Forschung und Materialwissenschaft (MoMaF), Literaturwissenschaft (SDC4Lit).

⁶Projektlaufzeit: 01.05.2023-31.07.2027; Vorgängerprojekt: bw2FDM (Projektlaufzeit: 01.05.2019-30.04.2023).

Gelingendes Forschungsdatenmanagement ist unter anderem in hohem Maße abhängig von einer optimalen Unterstützung durch Infrastrukturen für das Generieren und Analysieren sowie das Teilen und Speichern von Forschungsdaten. Bei den Kapazitäten für den Datenaustausch hat Baden-Württemberg ein deutschlandweit einzigartiges Instrument zur Unterstützung der Forschung und Lehre geschaffen: Das Wissenschaftsnetz *BelWü* verbindet ca. 100 wissenschaftliche Einrichtungen im Land. Die durch das Wissenschaftsnetz verbundene gemeinsame Infrastruktur umfasst dabei Cloud- und Filesharing-Dienste, serverbasierte Dienste sowie standortübergreifende Lösungen für datenintensive Simulationen und Experimente. Der Aufbau des Bereichs Hoch- und Höchstleistungsrechnen (High Performance Computing – HPC) und datenintensiver Dienste wurde seit 2008 vorangetrieben und damit die Möglichkeiten einer gemeinsamen, koordinierten Nutzung von Ressourcen ausgebaut. Seit 2012 wird dieser Ausbau durch das Umsetzungskonzept *bwHPC* sowie die darauf aufbauenden Rahmenkonzepte strategisch gestützt.

Herausforderungen und Perspektiven

Das frühe und breite Engagement Baden-Württembergs im Bereich des Forschungsdatenmanagements und die dadurch im Land aufgebaute Expertise sind jedoch nicht nachhaltig abgesichert: Die wenigsten innerhalb der o. g. Förderlinien aufgebauten Dienste und Strukturen im Forschungsdatenmanagement werden von Personal betreut, das auf Dauerstellen beschäftigt wird. Laufen Projektfördermittel aus oder wechselt das Projektpersonal, führt dies im schlechtesten Fall zur Einstellung von Diensten und dem Zerfall ressourcenintensiv aufgebauter, qualitativ hochwertiger Strukturen. Auch die Forschungsdaten selbst müssen nachhaltig gespeichert werden (können) – auch über eine Frist von zehn Jahren hinaus.

Am ehesten sind in Baden-Württemberg nachhaltige Strukturen im Forschungsdatenmanagement aktuell in der direkten Beratung und Schulung von Forschenden zu finden: Einschlägig qualifiziertes Personal – zumeist an wissenschaftlichen Bibliotheken und/oder Rechenzentren angesiedelt – berät und schult Forschende zum Forschungsdatenmanagement (z. B. zu FDM-Anforderungen von Drittmittelgebern, Lizenzwahl, Repositorienwahl, IT-Lösungen für spezielle Anforderungen). Während die baden-württembergischen Universitäten den Aufbau entsprechender Teams in den vergangenen Jahren vorangetrieben haben, folgen nun die Hochschulen für Angewandte Wissenschaften und andere Hochschularten. Auch fach-wissenschaftliche Expertise im Forschungsdatenmanagement (abgedeckt z. B. durch sog. Data Stewards oder Data Scientists) wird immer stärker nachgefragt. Zudem ist bereits jetzt ein Engpass an Fachkräften für die Betreuung von Hard- und Software für das Forschungsdatenmanagement zu beobachten.

In den vergangenen Jahren hat sich die Forschungsdatenlandschaft in Baden-Württemberg immer weiter ausdifferenziert. Im selben Maß müssen nun steuernde und regelnde Strukturen geschaffen werden, die dem Aufbau einer dauerhaften, flächendeckenden Unterstützung des Forschungsdatenmanagements dienen. Aktuelle datenrechtliche

Entwicklungen⁷ wirken sich auf diesen Strukturaufbau ebenso aus wie (hochschul)politische Entscheidungen. Eine auf Langfristigkeit ausgelegte Governance unterstützt den Wandel zu einer datenbasierten, digitalen Gesellschaft.

Auch auf technischer Seite müssen nachhaltige Systeme für das Forschungsdatenmanagement geschaffen werden: Auf Basis des Wissenschaftsnetzes *BelWü* werden im Rahmen des *bwDATA*-Konzepts gemeinsame langfristige Speicherlösungen für Forschungsdaten angestrebt. Diese sind jedoch im Vergleich zu den Kapazitäten für das Rechnen mit Forschungsdaten aktuell noch unterentwickelt. Insbesondere für die Langzeitdatenspeicherung von zehn und mehr Jahren gibt es noch kaum tragfähige und vor allem keine langfristig finanzierten Modelle. Eine Herausforderung ist es auch, verschiedene Systeme untereinander kompatibel zu gestalten. Erst wenn dies gelingt, ist ein Datenaustausch zwischen verschiedenen Einrichtungen und Ebenen problemlos und ohne größeren Nachbereitungsaufwand möglich. Daneben ist es notwendig, auch die ökologische Nachhaltigkeit technischer Systeme (Green IT) im Blick zu behalten.

Gelingendes Datenmanagement spielt bei der Bewältigung der allgegenwärtigen Digitalisierungsprozesse eine Schlüsselrolle. Es muss ein funktionierendes Datenökosystem entwickelt werden, das Daten verfügbar hält und zugänglich macht. Die baden-württembergische Forschungslandschaft kann hier mit ihrer innovativen Kraft vorangehen, neue Ideen und Modelle entwickeln, erproben und nachhaltig für sich und die Gesellschaft nutzbar machen. Hierfür bedarf es der strategischen Weiterentwicklung des Forschungsdatenmanagements und Investitionen in die Forschungsdateninfrastruktur, die sich in drei Handlungsfelder gliedern lassen:

1. Die Aus- und Weiterbildung von **Personal** sowohl in der Wissenschaft als auch an Infrastruktureinrichtungen sowie die Möglichkeit der Einstellung bereits ausgebildeter Expertinnen und Experten.
2. Die Entwicklung und Etablierung einer **Governance** im Hinblick auf geteilte Zuständigkeiten und Ressourcen.
3. Die Beschaffung und Instandhaltung sowie die Entwicklung **technischer Lösungen** für den Umgang mit Forschungsdaten.

⁷Etwa im Kontext der europäischen Digitalstrategie (https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age_de) (29.09.2023), der deutschen Digitalstrategie (<https://digitalstrategie-deutschland.de/>) (29.09.2023) bzw. (<https://www.bmv.de/DE/Themen/Digitales/Digitalpolitische-Strategien-und-Schwerpunkte/Digitalstrategie-der-Bundesregierung/digitalstrategie-der-bundesregierung.html>) (13.10.2025) oder der baden-württembergischen Digitalisierungsstrategie digital.LÄND (<https://im.baden-wuerttemberg.de/de/digitalisierung/digitalisierungsstrategie/strategie-im-ueberblick>) (29.09.2023).

Handlungsfeld Personen

Die in Baden-Württemberg gewonnenen Erfahrungen im Strukturaufbau zeigen, dass es für die praktische Umsetzung und konzeptionelle Weiterentwicklung des Forschungsdatenmanagements notwendig ist, verschiedene Personengruppen in den Prozess einzubeziehen bzw. zu adressieren: Wissenschaftlerinnen und Wissenschaftler, forschungsunterstützendes Personal an Infrastruktureinrichtungen (insbesondere z. B. Bibliotheken und Rechenzentren) und nicht zuletzt Studierende und Lehrende an den Hochschulen. Für alle genannten Gruppen müssen daher vorhandene Aus- und Weiterbildungsangebote sinnvoll vernetzt, gezielt weiterentwickelt und wo nötig ergänzt werden.

Ausgangslage und Perspektive

Wissenschaftlerinnen und Wissenschaftler

Der Umgang mit Daten gehört für Wissenschaftlerinnen und Wissenschaftler zum Forschungsalltag. Ob Forschungsdaten jedoch im Sinne eines nachhaltigen Forschungsdatenmanagements beschrieben und archiviert werden, hängt aktuell noch stark vom Wissensstand der Forschenden, der Verankerung des Forschungsdatenmanagements in der jeweiligen Fachcommunity und dem Wissen über Zugang und Nutzung lokaler (IT-)Infrastrukturen zum Forschungsdatenmanagement ab. Der Aufbau der NFDI sorgt mittelfristig für ein zunehmendes fachspezifisches Wissen zum Forschungsdatenmanagement und etabliert FDM-Standards für verschiedene Fachcommunities. Dadurch steigen jedoch gleichzeitig die Anforderungen an einzelne Forschende: Von Seiten der Förderinstitutionen und der Forschungseinrichtungen wird erwartet, dass Wissenschaftlerinnen und Wissenschaftler bereits im Vorfeld von Forschungsvorhaben z. B. Datenmanagementpläne erstellen und sich der Möglichkeiten und Herausforderungen bei der fachgerechten Speicherung sowie Veröffentlichung ihrer Daten bewusst sind. Dazu benötigen Forschende einen guten Überblick über die ihnen zur Verfügung stehende Infrastruktur, z. B. Repositorien, sowie allgemeine Beratungs- und Unterstützungsangebote. Für Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftler bieten sich hierfür systematische Einführungen ins Forschungsdatenmanagement an, die auch diese Aspekte abdecken. Etablierte Forscherinnen und Forscher benötigen dagegen vor allem einschlägige Unterstützungsangebote, die bei der Umsetzung fachlicher FDM-Standards mit Hilfe der am Standort vorhandenen Infrastruktur beraten.

Wissenschaftsunterstützendes Personal

Durch die zunehmende Bedeutung datenorientierter Wissenschaft wird der Bedarf an hervorragend ausgebildetem Personal, das lokal an Hochschulen und Forschungseinrichtungen zur Verfügung steht, stetig größer. Zwar finden sich insbesondere an Infrastruktureinrichtungen wie Rechenzentren und Bibliotheken bereits heute allgemeine Beratungs- und Informationsstellen, die Forschende beim Forschungsdatenmanagement unterstützen. Aber diese können dem steigenden Bedarf nur bedingt Rechnung tragen. Die wachsenden Ansprüche der Fachcommunities durch Etablierung fachspezifischer

FDM-Standards benötigen ebenfalls einschlägig qualifiziertes Personal. Wissenschaftliche Mitarbeiterinnen und Mitarbeiter mit FDM-Expertise finden sich vereinzelt etwa in größeren Forschungsprojekten oder an den Fakultäten, wo sie als direkte Ansprechpartner für FDM-Belange fungieren. Eine fachspezifische Unterstützung einzelner Forschender durch die NFDI wird gerade aufgebaut. Allerdings wird diese auf absehbare Zeit nicht in der Breite verfügbar sein, basiert auf Projektstrukturen, und ist ebenfalls darauf angewiesen, Personal mit entsprechenden Kompetenzen zu finden oder zu qualifizieren. Eine fachspezifische FDM-Unterstützung erhalten Wissenschaftlerinnen und Wissenschaftler, die nicht direkt in die Strukturen der NFDI eingebunden sind, daher bislang noch sehr selten. Durch die Schaffung entsprechender lokaler Stellen sowie den Aufbau von allgemeiner und fachspezifischer Expertise sollte diesem Mangel entgegengewirkt werden. Diese Stellen können zudem die Brücke zwischen der lokalen Fachcommunity und der NFDI bilden und so eine hohe Qualität der Unterstützung nahe an aktuellen Standards sicherstellen.

FDM-Fachpersonal benötigt eine entsprechende Aus- und Weiterbildung, um die Forschenden optimal bei der Digitalisierung von Forschungsprozessen unterstützen zu können. Notwendig sind Angebote, die eine große Bandbreite von Themen abdecken: Grundlegend ist Wissen über die vorhandenen Strukturen zur Unterstützung des Forschungsdatenmanagements an der eigenen Hochschule, im Bundesland und auf (inter-)nationaler Ebene (insbesondere NFDI) und der von diesen erbrachten Diensten. Darüber hinaus benötigt FDM-Personal auch Kenntnisse bezüglich Daten-Handling und Metadatenbeschreibungen bzw. -formaten ebenso wie zu verwandten Themen wie Rechtemanagement und Lizenzen. Zudem bedarf es darüber hinaus für die Betreuung von Großgeräten und Speichersystemen sowie der Beratung zu Forschungssoftware der Expertise hochspezialisierten IT-Personals.

Um kompetentes forschungsunterstützendes Personal im Bereich Forschungsdatenmanagement zu gewinnen, sollten entsprechende Karrierewege von der fachwissenschaftlichen Forschung in die Wissenschaftsunterstützung vorgezeichnet und geebnet werden. Damit eine Beschäftigung im forschungsunterstützenden Bereich von Nachwuchswissenschaftlerinnen und -wissenschaftlern in Betracht gezogen wird, müssen Berufsbilder (z. B. „Data Steward“) in ihrem Qualifikationsprofil und Aufgabengebiet geschärft werden. Außerdem muss mit der entsprechenden Expertise eine angemessene Entlohnung einhergehen.

Studierende und Lehrende

Kenntnisse im Forschungsdatenmanagement sind nicht nur für Forschende und wissenschaftsunterstützendes Personal essentiell. Sowohl in den Naturwissenschaften als auch in den Geistes- und Sozialwissenschaften müssten entsprechende fachspezifische Kenntnisse selbstverständlicher Teil des Lehrinhalts sein. Bedingt durch unterschiedliche Fachkulturen und -anforderungen sowie den insgesamt sehr unterschiedlichen Aufbau von Studiengängen gibt es jedoch bislang noch keine einheitlichen Mindeststandards für Inhalte des Forschungsdatenmanagements in der Lehre bzw. werden entsprechende Inhalte nicht unter dieser Bezeichnung zusammengeführt. Hinzu kommt der unterschiedliche Kenntnisstand zum Forschungsdatenmanagement bei Hochschullehrerinnen und -lehrern, der einer Vereinheitlichung von Lehrinhalten ebenfalls entgegensteht. Grundlegend

und in Übereinstimmung mit der guten wissenschaftlichen Praxis sind Studierende insbesondere auf die Rolle des Forschungsdatenmanagements bei der Plagiatsvermeidung und Ergebnisreproduktion zu sensibilisieren. Es ist daher anzustreben, Aspekte guten Forschungsdatenmanagements flächendeckend in fachspezifische Methodeneinführungen zu integrieren. Lehrenden kann diese Aufgabe durch die Bereitstellung von Open Educational Resources (OER) erleichtert werden, die entsprechend angepasst und in die eigenen Lehrveranstaltungen integriert werden können. Fachspezifische Informationen für die Ausbildung von Studierenden werden auch von einigen NFDI-Konsortien bzw. übergreifend in der NFDI-Sektion Training & Education entwickelt.

Organisation

Forschungsdatenmanagement entwickelt sich nach wie vor sehr dynamisch. Daraus resultiert ein hoher und tendenziell steigender Ressourcen- und Personalbedarf. Um Doppelstrukturen zu vermeiden und effiziente Verfahren zu etablieren, ist es notwendig, perspektivisch eine Verständigung auf Regelungen und Leitlinien herbeizuführen. Es soll deshalb abgestimmt und etabliert werden, welche Dienste und Angebote auf welcher Ebene (Einrichtung, Bundesland, Bund) erbracht bzw. angesiedelt werden sollten. Dafür werden hier einige Grundprinzipien definiert, die im Handlungsfeld Governance intensiver diskutiert bzw. ausgeführt werden.

Ebene der Einrichtung

Jede Einrichtung sollte lokale bzw. institutionelle Informations-, Beratungs-, und Schulungsangebote für ihre Mitglieder zur Verfügung stellen und muss dafür das nötige Personal vorhalten. Idealerweise werden diese Angebote in abgestimmter Kooperation zwischen Infrastruktureinrichtungen und Fachbereichen erbracht. Die konkreten organisatorischen Ausprägungen können divergieren: Denkbar sind Modelle, die von einer Zusammenführung des Personals in einem zentralen Kompetenz- oder Servicezentrum bis hin zu dezentralen Organisationsformen reichen, in denen verschiedene FDM-Teams als koordinierter Verbund agieren. Um die passende Organisationsform für FDM-Unterstützungsangebote für die jeweilige Einrichtung zu bestimmen, ist ein strategischer Prozess notwendig, der die zentralen Akteure der Einrichtung miteinbezieht, die jeweiligen Bedarfe systematisch erfasst und eine Akzeptanz auf breiter Basis schafft.⁸

Forschungsdatenmanagement ist eine Langfristaufgabe, die aktuell noch sehr häufig durch Personen wahrgenommen werden muss, die nur über Projektstellen beschäftigt sind. Die aufgebaute Expertise muss für die jeweilige Institution jedoch nachhaltig nutzbar

⁸Ein Beispiel für eine lokale, durchdachte und weit in der Einrichtung ausgerollte Struktur bietet die TU Delft mit einer universitätsweiten Research Data Framework Policy, konkretisiert in Research Data Management Policies an jeder Fakultät sowie einer Koordinierungsstelle bestehend aus „Data Stewards“ mit Lotsen- und Brückenfunktion an jeder Fakultät und definierten Rollen von IT, Verwaltung etc. Vgl. hierzu: Ahlers, Ingeborg, *et al.* (2020): TU Delft Research Data Framework Policy (Version V2). Zenodo. <https://doi.org/10.5281/zenodo.4088123>; sowie TU Delft & Faculty Policies: <https://www.tudelft.nl/en/library/research-data-management/r/policies/tu-delft-faculty-policies> (29.09.2023).

gemacht werden. Es ist deshalb wichtig, dauerhafte Teamstrukturen zu etablieren und qualifiziertes Personal über die reinen Projektphasen hinaus an den Einrichtungen zu halten. Dies wird nur durch die Schaffung von Dauerstellen gelingen. Eine Anbindung von FDM-Aufgaben an Personen, deren sonstige Aufgaben von vornherein nicht auf Dauer angelegt sind, kann aus dem gleichen Grund allenfalls eine Ausnahme sein.⁹

Die Vermittlung von Methoden des Forschungsdatenmanagements in ihrer jeweiligen fachlichen Ausgestaltung ist ein wichtiger Baustein für die Ausbildung von Studierenden. Die Hochschulen sollten entsprechende Ausbildungsanteile daher in ihren Curricula bzw. Modulhandbüchern verpflichtend festschreiben. Diese fachspezifischen Angebote könnten durch Lehrformate sinnvoll ergänzt werden, die über die Infrastruktureinrichtungen angeboten werden, z. B. zu Publikationsformen von Daten, Metadatenbeschreibungen, der Verwendung von persistenten Identifikatoren, Langzeitarchivierung oder Rechtsfragen (insb. Datenschutz- und Datennutzungsrechte).

Regionale und nationale Ebene

Insbesondere die Aus- und Weiterbildung von wissenschaftsunterstützendem Personal sollte im Land Baden-Württemberg vorangetrieben und koordiniert werden. Hierzu müssen Angebote entwickelt und für alle Hochschulen niedrigschwellig zugänglich gemacht werden. Die Anforderungen bzw. Inhalte sollten vergleichbar und transparent gestaltet werden. Dabei ist darauf zu achten, dass die Angebote eine hohe Anschlussfähigkeit zu den Entwicklungen in anderen Bundesländern und der NFDI behalten. Die Koordination entsprechender Angebote wird, dem Vorbild anderer Bundesländer folgend¹⁰, die Landesinitiative *bwFDM* übernehmen. Karrierewege im Bereich Forschungsdatenmanagement könnten durch das Angebot entsprechender Studiengänge oder Aufbaumodule aufgezeigt werden. Die Einrichtung solcher Ausbildungsangebote sollte geprüft werden.

Sehr wichtig ist auch die Vernetzung der in diesem Bereich tätigen Personen, z. B. „Data Stewards“ und Forschenden, um im aktiven Austausch voneinander zu lernen und Ideen zu entwickeln. Das Projekt *bw2FDM*¹¹ hat die Vernetzung von FDM-Expertinnen und -Experten auf Landesebene begleitet, die Landesinitiative *bwFDM* wird diese Bemühungen mit der Bereitstellung passender zentraler Strukturen weiter ausbauen. Eine optimale Aufstellung von Vernetzungsstrukturen, die ein nachhaltiges Forschungsdatenmanagement unterstützen, bedarf der Abstimmung mit allen Akteuren.

⁹Dies deckt sich auch mit den Forderungen des Koalitionsvertrags der Bundesregierung 2021-2025, S. 23: <https://www.bundesregierung.de/resource/blob/974430/1990812/1f422c60505b6a88f8f3b3b5b8720bd4/2021-12-10-koav2021-data.pdf?download=1> (29.09.2023) bzw. https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag_2021-2025.pdf (13.10.2025).

¹⁰Nordrhein-Westfalen: Zertifikatskurs Forschungsdatenmanagement der TH Köln in Zusammenarbeit mit der FDM-Bundeslandinitiative *fdm.nrw* und ZB MED (Informationszentrum Lebenswissenschaften), https://www.th-koeln.de/weiterbildung/zertifikatskurs-forschungsdatenmanagement_82048.php (29.09.2023).

¹¹Baden-württembergisches Begleit- und Weiterentwicklungsprojekt für Forschungsdatenmanagement: <https://bwfdm.de/ueber-uns/vorgaengerprojekte/> (29.09.2023).

Um zu bestmöglicher Abstimmung zu kommen und Synergieeffekte zu ermöglichen, sollten die lokalen Einrichtungen mit der Landesebene und der NFDI eng zusammenarbeiten. Gerade komplementäre Angebote können helfen, Doppelstrukturen zu vermeiden und den Austausch zu fördern. So könnten zum Beispiel über NFDI-Konsortien oder Fachinformationsdienste Kompetenzen vorgehalten werden, die fachspezifisch und überregional unterstützen, aber lokal angesiedelt sind.

Es ist in diesem Zusammenhang darauf hinzuweisen, dass viele Infrastruktur-Mitarbeitende und Forschende die Vernetzung bislang „nebenher“ vorangetrieben haben, d. h. zusätzlich zu ihren eigentlichen Aufgaben. Durch die gestiegene Komplexität der Strukturen wird dieses Vorgehen zunehmend zur Belastung für die involvierten Personen. Es gilt daher, schlanke Strukturen zu schaffen, die nicht zur Überforderung der beteiligten Mitarbeitenden führen. Gleichzeitig sind Politik und Hochschulen gleichermaßen aufgefordert, die unbestreitbar notwendige strategische Vernetzung durch einen entsprechenden Einsatz personeller Ressourcen zu unterstützen.

Empfehlungen

- Die optimale Unterstützung von Wissenschaftlerinnen und Wissenschaftlern, FDM-Personal, Studierenden und Lehrenden kann nur im Dialog gestaltet werden. Die konkrete Ausprägung und Umsetzung von Informationsangeboten sowie Aus- und Weiterbildungsstrukturen müssen daher unter Beteiligung der wissenschaftlichen Einrichtungen, beteiligten Akteure und Stakeholder erfolgen. Dabei kann der Zukunftsworkshop „Hochschulen in der digitalen Welt“ aus dem Jahr 2022 exemplarisch als erfolgreiches Format dienen.
- Die Entwicklung entsprechender Qualifikations- und Professionalisierungsformate bis hin zu Zertifikaten für wissenschaftsunterstützendes Personal sollte kooperativ erfolgen, um Zielerreichung, Effizienz und Qualitätssicherung zu gewährleisten.
- Für den wissenschaftlichen Nachwuchs ebenso wie für etablierte Wissenschaftlerinnen und Wissenschaftler müssen an allen Hochschulen Beratungs- und Unterstützungsangebote zur Datenkompetenz aufgebaut werden. Ein Austausch lokaler Einrichtungen mit baden-württembergischen Strukturen und der NFDI ist in dieser Hinsicht anzustreben, um komplementäre Angebote zu schaffen.
- Die Berücksichtigung des FDM-Bedarfs an Hochschulen für Angewandte Wissenschaften und anderen Hochschularten sowie die Entwicklung passender Angebote ist notwendig. Auch hier gibt es im Land bereits Vorbilder (z. B. bei der Implementierung von Open-Access-Beratungsstrukturen oder beim Zentralen OER-Repositorym ZOERR).
- Es müssen Karrierewege für FDM-Personal etabliert und geebnet werden. Die Einrichtung entsprechender Studiengänge sollte geprüft werden.

Handlungsfeld Governance

Die Forschungsdatenlandschaft und das Forschungsdatenmanagement in Baden-Württemberg unterliegen steuernden und regelnden Strukturen, die auf unterschiedlichen Ebenen angesiedelt sind. Um eine breite Akzeptanz und Partizipation aller Akteure aus Wissenschaft, Wissenschaftsinfrastruktur, Wirtschaft und Politik zu erreichen, müssen neue Governance-Konzepte entworfen und mit vorhandenen verknüpft werden. Dadurch werden Zuständigkeiten und Entscheidungsstrukturen transparent, was eine wesentliche Voraussetzung darstellt, um die Auffindbarkeit, den Zugang, die Speicherung und die Nachnutzbarkeit von Forschungsdaten nachhaltig zu sichern.

Gelungende Governance ermöglicht eine starke Vernetzung und vertrauensvolle Zusammenarbeit zwischen den beteiligten Akteuren und verhindert damit die Entstehung von Konflikten. Zugleich unterstützt sie einen effizienten Einsatz von Ressourcen, da der Aufbau von Doppelstrukturen durch eine klare Aufgabenteilung weitestgehend vermieden werden kann. Über Governance-Konzepte können darüber hinaus Diskurse organisiert werden, die eine rechtssichere und ethisch fundierte Erzeugung, Verarbeitung und Nachnutzung von Forschungsdaten erst ermöglichen. Die Forschungsdateninfrastruktur in Baden-Württemberg ist vielfältig und ihre Dienste und Angebote werden auf unterschiedlichen Ebenen erbracht – vor Ort in den jeweiligen Einrichtungen, auf Landesebene, aber auch auf Bundesebene und darüber hinaus. Die Vernetzung dieser notwendigen Akteure sowie die koordinierte Steuerung ihrer Aktivitäten ist noch defizitär. Viele strategisch bedeutsame Bereiche weisen außerdem Leerstellen auf, an denen Akteure mit einschlägiger Expertise bislang noch fehlen. Die vorhandenen Strukturen müssen daher intelligent vernetzt, ergänzt und strategisch gefördert werden, um allen Hochschulen und Forschungseinrichtungen Zugang zu einer adäquaten Forschungsdateninfrastruktur in Form von personellen Ressourcen und IT-Lösungen zu bieten.

Ausgangslage und Perspektive

Organisation forschungsunterstützender Infrastruktur

Die Anforderungen an Forschende bezüglich des Umgangs mit Forschungsdaten sind durch die Vorgaben von Drittmittelgebern und Hochschulen immer weiter gestiegen. Alle Hochschulen in Baden-Württemberg brauchen Anlaufstellen, die Forschende beraten und gegebenenfalls an Expertinnen und Experten mit spezialisiertem Wissen weiterleiten können. In den allermeisten Fällen gibt es derartige Anlaufstellen bereits. Deren Anbindung an die NFDI ist in diesem Zusammenhang besonders wichtig, um die fachliche Expertise der Konsortien, die bislang aufgrund fehlender Strukturen meist nur für eng in die NFDI-Strukturen eingebundene Einrichtungen verfügbar ist, auch für andere Standorte verlässlich nutzbar zu machen. Hierfür benötigt das wissenschaftsunterstützende Personal Strukturen, die eine Vernetzung der Einrichtungen untereinander und zur NFDI ebenso sicherstellen, wie den Zugang zu einer einschlägigen, qualitativ hochwertigen Aus- und Weiterbildung für FDM-Personal und Forschende. Um die Bedarfe aller Einrichtungen im Blick zu behalten und daran ausgerichtete Verbesserungsmaßnahmen für die Unterstützung

von Forschenden und wissenschaftsunterstützendem Personal voranzutreiben, bedarf es der Verstärkung der im Projekt bwFDM aufgebauten Koordinierungs- und Kontaktstelle, bei der die entsprechenden Informationen zusammenlaufen.

Auf der Ebene der technischen und infrastrukturellen Unterstützung des Forschungsdatenmanagements hat es sich bewährt, dass größere Universitäten und Forschungseinrichtungen bestimmte Infrastrukturen und Teilprozesse des Forschungsdatenmanagements passgenau auf ihre jeweiligen institutionellen Bedarfe zuschneiden bzw. diese entwickeln. Hierbei gilt es jedoch, darauf zu achten, dass die eingesetzten Systeme und Infrastrukturen interoperabel sind und damit nicht nur innerhalb einer Einrichtung, sondern auch standortübergreifend kooperative Forschungsprozesse unterstützen. Insbesondere für kleinere Universitäten, Hochschulen und Forschungseinrichtungen ist es hingegen häufig aus organisatorischen und finanziellen Gründen nicht sinnvoll oder nicht möglich, eigene Infrastrukturen für die Unterstützung des Forschungsdatenmanagements zu schaffen. Deshalb sollten einrichtungsübergreifende Strukturen entstehen, die lokale Stellen, wo angemessen, ergänzen und entlasten.

Insgesamt gilt, dass der Aufbau forschungsunterstützender Infrastruktur weiterhin noch zu häufig in Form von befristeten Projekten erfolgt. Die Wahrnehmung von Beratungsaufgaben im Bereich Forschungsdatenmanagement, die Weiterentwicklung von forschungsunterstützenden FDM-Angeboten sowie die Bereitstellung von IT-Infrastruktur, insbesondere für die Langzeitarchivierung, sind jedoch Daueraufgaben. Es muss daher sichergestellt werden, dass diese langfristigen Aufgaben auch dauerhaft und verlässlich finanziert werden.

Strategische Vernetzung

Der Prozess der Vernetzung zwischen den baden-württembergischen Forschungseinrichtungen ist auf die strategische Unterstützung durch das Land Baden-Württemberg angewiesen. Die vier Science Data Center, die erfolgreich an die NFDI angebunden wurden, sind ein sehr gutes Beispiel, wie sich die lokale Vernetzung durch entsprechende Ausschreibungen gezielt fördern lässt; auch an die guten Erfahrungen aus den *bwHPC*-Projekten kann in dieser Hinsicht angeknüpft werden. Andere Vernetzungsgremien wie der Arbeitskreis Forschungsdatenmanagement (*AK FDM*), der als Austauschplattform der baden-württembergischen Universitäten auf Arbeitsebene dient, haben sich in Bottom-Up-Ansätzen aus dem Bedarf nach regelmäßiger gegenseitiger Information und der Suche nach Kooperationsmöglichkeiten gegründet. Es ist wünschenswert, die bislang aufgebauten Strukturen untereinander noch stärker zu vernetzen. Zusätzlich gilt es selbstverständlich, bislang noch nicht in einschlägigen Gremien vertretene Hochschulen ebenfalls in die Vernetzung einzubeziehen.

Die Vernetzung der Landesebene zu regionalen (z. B. FDM-Landesinitiativen), nationalen (z. B. NFDI-Konsortien) und internationalen Akteuren (z. B. *EOSC*, *GO FAIR* etc.) ist noch ausbaufähig. In vielen Fällen besteht sie derzeit oft nur aufgrund einer mehr oder minder zufälligen Personalunion in unterschiedlichen Gremien. Baden-Württemberg ist in einer besonders starken Position, um die Vernetzung der verschiedenen Ebenen

strategisch voranzutreiben: Bereits etabliert sind die Plattform forschungsdaten.info¹² sowie die *E-Science-Tage*¹³ als deutschlandweit führendes Veranstaltungsformat. Darüber hinaus stehen baden-württembergische Forschungseinrichtungen in stetigem Austausch untereinander und sind intensiv an der NFDI beteiligt.¹⁴ Dieser derzeitige Vorsprung im Ländervergleich sollte genutzt werden, um den Forschungsstandort Baden-Württemberg dauerhaft weiter zu stärken. Gleichzeitig ist zu beobachten, dass der Bereich Vernetzung in anderen Bundesländern in letzter Zeit – auch finanziell – massiv gestärkt wurde. Das Land Baden-Württemberg wird daher Anstrengungen unternehmen, um den Spitzenplatz im Forschungsdatenmanagement zu behaupten und auszubauen.

Informationen zur Praxis des Forschungsdatenmanagements sowie Vorschläge zur Weiterentwicklung müssen noch stärker als bislang auf die politische Entscheidungsebene getragen werden. Zwar gibt es bereits eine Reihe von stärker operativ arbeitenden Gremien (*AK FDM*, *AG BibDir*, *ALWR*), die jedoch gerade wegen ihres operativen Charakters für übergreifende Fragen der strategischen Ausrichtung nur bedingt geeignet sind. Der Landesnutzerausschuss für Forschungsinfrastrukturen (LNA) als wissenschaftliches Beratungsgremium beschäftigte sich bisher vorwiegend mit dem High Performance Computing (HPC) und Data Intensive Computing (DIC). Es sollte deshalb geprüft werden, ob die genannten Strukturen von einem Landesdatenrat – bestehend aus ausgewiesenen Expertinnen und Experten – sinnvoll ergänzt werden können und sollten. Ein solches Expertengremium könnte auch das Ministerium bei Bedarf zu Fragen zum Umgang mit Forschungsdaten beraten und wäre geeignet, die Umsetzung und Fortentwicklung der Forschungsdatenstrategie zu begleiten.

Recht und Datenschutz

In allen Bereichen der Digitalisierung kann eine sehr dynamische Entwicklung des Rechts beobachtet werden. Insbesondere die Bereiche Datenschutzrecht, Urheberrecht und Nutzungsrechte sind für das Forschungsdatenmanagement von besonderer Bedeutung. Die Geschwindigkeit der Entwicklung macht es jedoch zunehmend schwieriger, Gesetzgebungsprozesse im Sinne eines möglichst offenen und FAIRen Forschungsdatenmanagements zu begleiten und zu unterstützen. Es müssten daher Strukturen aufgebaut werden, die eine deutlich stärkere Beratung zu und ein Einbringen von eigenen Belangen in forschungsdatenrelevante Gesetzgebungsprozesse auf europäischer und nationaler Ebene ermöglichen, als es derzeit der Fall ist (Data Act¹⁵, Artificial Intelligence Act¹⁶, Forschungsdatengesetz¹⁷ etc.).

¹²<https://forschungsdaten.info/> (29.09.2023).

¹³<https://e-science-tage.de/> (29.09.2023).

¹⁴Begleitet werden sie hierbei durch die Landesinitiative bwFDM.

¹⁵<https://digital-strategy.ec.europa.eu/en/policies/data-act> (29.09.2023).

¹⁶<https://artificialintelligenceact.eu/> (29.09.2023).

¹⁷<https://www.bmbf.de/bmbf/shareddocs/kurzmeldungen/de/2023/03/230306-forschungsdatengesetz.html> (29.09.2023) bzw. <https://www.bmftr.bund.de/DE/Service/Gesetze/Gesetze-Einzel/Forschungsdatengesetz.html?templateQueryString=forschungsdatengesetz> (13.10.2025).

Die gute wissenschaftliche Praxis fordert von Hochschulen und Forschungseinrichtungen einen sorgfältigen Umgang mit Forschungsdaten, der über reine Rechtskonformität hinausgeht. Die gemeinsame Erarbeitung und Weiterentwicklung von entsprechenden Policies und anderen Hinweisen zum Forschungsdatenmanagement über alle baden-württembergischen Forschungseinrichtungen hinweg ist wünschenswert, um auch hier die Interoperabilität sicherzustellen und einrichtungsübergreifende Forschungsprozesse zu erleichtern.

Vonseiten der Wissenschaftlerinnen und Wissenschaftler besteht ein anhaltend hoher Bedarf an Beratung zu rechtlichen FDM-Themen. Dies betrifft nicht nur datenschutzrechtliche Fragen, deren Beantwortung aufgrund neuer technischer Entwicklungen zunehmend komplexer wird. Vielmehr benötigt es auch in der Beratung zu anderen datenrechtlichen Fragen wie Urheber- und Nutzungsrechten spezialisiertes Wissen, das durch die generalistisch aufgestellten Justizariate der Hochschulen und Forschungseinrichtungen nicht abgedeckt werden kann. Eine juristische Begleitung von FDM-Prozessen unter Berücksichtigung der dynamischen Entwicklung des Rechts ist jedoch essentiell, um unter sich wandelnden Bedingungen rechtssichere Spitzenforschung an und mit Daten betreiben zu können. Daher sollte die Einrichtung einer zentralen Stelle mit der entsprechenden Kompetenz, beispielsweise am Justizariat einer Hochschule, geprüft werden. Eine solche Stelle könnte auch das zuständige Ministerium in entsprechenden Fragen beraten.

Grundsätzlich ist es das Ziel, dass Forschungsdaten im Sinne von Open Science nach dem Grundsatz „so offen und nachnutzbar wie möglich“ gespeichert und zugänglich gemacht werden. Eine gemeinsame Praxis aller baden-württembergischen Hochschulen, die eine Offenheit von Forschungsdaten als Regelfall (*open by default*) festlegt, ist erstrebenswert.

Gleichzeitig gilt es, sensible Daten zu schützen, die beispielsweise personenbezogene Informationen enthalten oder Geschäftsgeheimnisse von Partnern aus der Wirtschaft berühren. Die Etablierung von Datentreuhandstellen¹⁸ könnte in diesen Fällen einen wesentlichen Beitrag zur Nachnutzung von Daten leisten. Entsprechende Pilotprojekte zeigen, dass dafür nicht nur Organisationsstrukturen und IT-Infrastrukturen aufgebaut werden müssen. Vielmehr ist es auch nötig, das Vertrauen der beteiligten Akteure zu gewinnen. Voraussetzung hierfür ist die Schaffung verbindlicher Regelungen im Bundes- und Landesrecht durch die Konkretisierung des auf europäischer Ebene bereits angelegten Rechtsrahmens für Datentreuhandstellen. Landeseigene Datentreuhandstellen können einen wesentlichen Beitrag für die Datensouveränität des Innovationsstandorts Baden-Württemberg leisten.

Ethik

Ein verantwortungsbewusster, von ethischen Grundsätzen geleiteter Umgang mit Forschungsdaten ist Teil der guten wissenschaftlichen Praxis. In allen Phasen des Forschungsdatenlebenszyklus, insbesondere jedoch bei der Erhebung, Veröffentlichung und Nachnutzung von Forschungsdaten, müssen die Auswirkungen von Forschungsar-

¹⁸Zur möglichen Ausgestaltung von Datentreuhandstellen vgl.: Rat für Informationsinfrastrukturen: Stellungnahme Datentreuhandstellen gestalten – Zu Erfahrungen der Wissenschaft, Göttingen 2020, <https://rfii.de/?p=4259>.

beiten auf Mensch und Umwelt berücksichtigt werden. Datenethische Fragen gehen dabei weit über daten(schutz)rechtliche Belange und die Forderung nach FAIRem Forschungsdatenmanagement hinaus. Sie umfassen unter anderem die korrekte Zuordnung von wissenschaftlicher Leistung zu einzelnen Forschenden oder den ethischen Umgang mit Versuchspersonen und Versuchstieren. Zudem spielen die Umweltrelevanz und die Abschätzung des Gefahrenpotenzials von Forschung (z. B. durch Unfälle und Missbrauch von Forschungsergebnissen) eine Rolle. Nicht zuletzt vor dem Hintergrund aktueller (sicherheits)politischer Entwicklungen müssen sich Forschende verstärkt Fragen nach dem Austausch von Wissen und Technologie mit Drittländern sowie dem Umgang mit sicherheitsrelevanter Forschung (*Dual Use*) stellen. Vielfach existieren für die genannten Themenkomplexe bereits rechtliche Regelungen; dennoch gibt es immer wieder Bereiche, in denen die Erhebung, Veröffentlichung oder Nachnutzung von Forschungsdaten auch aus ethischen Gründen problematisch ist.

Datenethische Fragestellungen sind für ein gutes Forschungsdatenmanagement vor allem bei der Frage nach der Zugänglichmachung von Daten enorm wichtig, daher ist der wissenschaftliche und gesellschaftliche Diskurs hierzu unbedingt zu fördern. Daneben bedarf es gewisser verbindlicher Regelungen für den Umgang mit ethischen Grenzfällen. Ein Beispiel für die Umsetzung von ethischen Überlegungen im Bereich Forschungsdatenmanagement in ein konkretes Regelwerk sind die CARE-Prinzipien für indigene Data Governance¹⁹, die Fragen der Kontrolle indigener Gemeinschaften über Anwendung und Nutzung von Daten behandeln, die über die jeweilige indigene Gemeinschaft oder auf ihrem Gebiet erhoben wurden.

Aufgrund der Komplexität des Gegenstands bedarf es auch im Bereich der Datenethik Stellen zur Unterstützung von Wissenschaftlerinnen und Wissenschaftlern, die sie bei Fragen juristisch beraten oder auf ethische Problematiken in Forschungsvorhaben hinweisen können. Dabei gilt es zu berücksichtigen, dass mittlerweile immer mehr Forschungsförderer und Wissenschaftsverlage Nachweise zur ethischen Unbedenklichkeit von Forschungsvorhaben fordern. Daher sollten an allen Forschungseinrichtungen entsprechende Beratungs- und Prüfstellen eingerichtet und auch auf deren angemessene personelle Ausstattung geachtet werden.

Auch ein ethischer Umgang mit den persönlichen Daten der Forscherinnen und Forschern ist nötig: Die Aggregation und Verwendung von persönlichen Daten, die beim Forschungsprozess entstehen, wird immer stärker von großen Wissenschaftsverlagen monetarisiert. Insbesondere die Deutsche Forschungsgemeinschaft (DFG) hat darauf hingewiesen, dass kommerzielles Datentracking hohe Gefahren für die Transparenz und Unabhängigkeit wissenschaftlicher Forschung birgt.²⁰

Datensouveränität

Selbst ein grundsätzlich so offen wie möglich entwickeltes Wissenschaftssystem darf nicht gänzlich unkontrolliert sein und auch eine Freigabe von Daten als FAIR zur Weiternutzung in anderen Forschungsprojekten darf nicht dazu führen, dass Forscherinnen und Forscher

¹⁹Global Indigenous Data Alliance: <https://www.gida-global.org/care> (29.09.2023).

²⁰DFG – Datentracking in der Wissenschaft, 2021, <https://doi.org/10.5281/zenodo.5900758>.

oder ihre Einrichtungen jede Verantwortung für und Kontrolle über ihre Daten verlieren. Dies gilt insbesondere für die Verwertung bzw. Vermarktung von Forschungsdaten durch Dritte.²¹ Die Missbrauchsgefahr durch einen unkontrollierten Datenabfluss oder einer Manipulation der Daten sind hoch; das betrifft nicht nur, aber insbesondere, den Schutz persönlicher Daten aller an der Forschung beteiligter Personen. Das allgemeine Thema Datensouveränität ist daher von großer Bedeutung auch im Bereich der Forschungsdaten. Die derzeitigen Schutzmechanismen zur Sicherung der Souveränität müssen kontinuierlich angepasst und weiterentwickelt werden.

Bewertung von wissenschaftlicher Leistung

Es besteht aus Forschung und Politik ein großes Interesse daran, sowohl wissenschaftliche Leistungen besser mess- und vergleichbar zu machen, als auch gesellschaftlichen Wert von Forschung adäquat abzubilden. Die derzeit noch vorherrschende Bewertung von Forschungsleistung überwiegend nach starren Publikations- und Zitationsindizes steht bereits seit längerem in der Kritik. Europaweit und international wird daher nach neuen Maßstäben für die Bewertung von Forschung, Forschenden und Forschungseinrichtungen gesucht, etwa in den Initiativen DORA²², CoARA²³ und ERA²⁴. Für Deutschland sollten diese neuen Bewertungssysteme im Einklang mit der DFG und dem Wissenschaftsrat entwickelt werden. Es bedarf eines wissenschaftlichen Kulturwandels²⁵, bei dem die wissenschaftliche Leistung nicht nur an klassischen Publikationsformen gemessen wird.

Ein stärkerer Einbezug von Forschungsdaten sollte ein wesentliches Element von neuen Bewertungssystemen sein. Aktuell werden Forschungsdaten noch nicht überall als Leistung neben Aufsätzen, Patenten und Monografien anerkannt und entsprechend zitiert. Eine neue Datenkultur sollte es mit sich bringen, dass die wissenschaftlichen Leistungen bei Erhebung und Verwaltung von Forschungsdaten höhere Anerkennung als derzeit finden. Das Land Baden-Württemberg wird die Fachcommunities bei der Entwicklung und Gestaltung neuer Bewertungssysteme für wissenschaftliche Leistung unterstützen.

Empfehlungen

- Das Projekt *bwFDM* (Landesinitiative Forschungsdatenmanagement Baden-Württemberg) dient den FDM-Akteuren als Koordinierungs- und Kontaktstelle. Im Rahmen einer nachhaltigen Governance sollte eine Verstetigung von *bwFDM* in Betracht gezogen werden.

²¹Vgl. Diskussionsimpuls des Rats für Informationsinfrastrukturen: Anforderungen an die Ausgestaltung eines Forschungsdatengesetzes und eines Dateninstituts, Juni 2023, <https://rfii.de/?p=9101>.

²²<https://sfdora.org/> (29.09.2023).

²³<https://coara.eu/> (29.09.2023).

²⁴https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/european-research-area_en (29.09.2023).

²⁵https://www.dfg.de/foerderung/info_wissenschaft/2022/info_wissenschaft_22_61/index.html (29.09.2023).

- Expertise für die juristische Beratung in datenrechtlichen Spezialfragen sollte zentral, z. B. an einem Justizariat einer Hochschule, angesiedelt werden.
- Für Fragen der Datenethik (auch über den Forschungskontext hinaus) sollte eine zentrale Beratungs- und Kontaktstelle an einer Hochschule aufgebaut werden.
- Datensouveränität bei Forschungsdaten muss als Thema adressiert werden. Ein stärkerer gesetzgeberischer Schutz sollte erwogen werden.
- Die Einrichtung eines Landesdatenrats von unabhängigen Expertinnen und Experten sollte geprüft werden.
- Datentreuhandstellen sollten gezielt aufgebaut und gefördert werden. Hierfür sollten entsprechende rechtliche Regelungen auf Landesebene geschaffen werden.
- Die Fortführung der interdisziplinären *E-Science-Tage* und der Informationsplattform *forschungsdaten.info* als zwei deutschlandweit wahrgenommene Austauschformate ist sicherzustellen.
- Dauerhafte Kosten, die für eine stabile Forschungsdateninfrastruktur unumgänglich sind, müssen auch dauerhaft finanziert werden. Entsprechende Geschäftsmodelle sind zu entwickeln.

Handlungsfeld Technik

Die Digitalisierung von Forschungsprozessen durchdringt alle Ebenen wissenschaftlicher Aktivität und setzt voraus, dass eine exzellente Infrastruktur und damit einhergehend auch Dienste für die Forschung bereitgestellt werden. Das dynamische und innovative Umfeld der Forschung erfordert stets aktuelle digitale Lösungen. Dies gestaltet sich besonders im Zusammenhang mit Forschungsdaten ausgesprochen herausfordernd: Die zu bearbeitenden Forschungsdaten steigen in manchen Anwendungsbereichen nahezu exponentiell an. Verarbeitungsprozesse von primären Daten hin zu wissenschaftlichen Erkenntnissen sind hochkomplex und müssen digital abgebildet werden. Die Entwicklung wird durch intelligente und automatisierte Ansätze, z. B. maschinelles Lernen und KI, maßgeblich verstärkt. Die baden-württembergische Wissenschafts-IT-Infrastruktur und ihre Dienste müssen auf diese Herausforderungen reagieren. Eine auf die Zukunft gerichtete Strategie ist Grundbedingung für die Fortführung der exzellenten Wissenschaft und Forschung am Standort Baden-Württemberg.

Schwerpunktmäßig ist die Bereitstellung einer zuverlässigen und leistungsfähigen IT-Infrastruktur, die eine verlustfreie Übertragung und Speicherung von Forschungsdaten über den gesamten Forschungsdatenlebenszyklus hinweg ermöglicht, eine *conditio sine qua non* für die erfolgreiche Durchführung von Forschungsaktivitäten. Auf dieser Grundlage können verlässliche IT-Services aufbauen, die eine benutzerfreundliche Arbeit an und mit Forschungsdaten sowie Forschungssoftware gewährleisten. Darüber hinaus bilden die IT-Infrastruktur und IT-Dienste auch die Voraussetzung für die Bereitstellung von

Software-Lösungen, die Forschende bei der Erhebung, Verwaltung und Kuratierung von Forschungsdaten unterstützen. Durch die anvisierte strategische Ausrichtung im Bereich des Forschungsdatenmanagements sollen wichtige Rahmenbedingungen geschaffen werden, die es Forscherinnen und Forschern im Land ermöglichen, reibungslos und IT-compliant aus Forschungsdaten wissenschaftliche Erkenntnisse zu erlangen.

Ausgangslage und Perspektive

Föderative Ansätze

In etlichen Forschungscommunities kann die Komplexität der zugrundeliegenden Fragestellungen nur über Forschungsverbünde bewältigt werden, die mehrheitlich interdisziplinär ausgelegt sind. Der Umgang mit Forschungsdaten in diesem Kontext erfordert föderative Ansätze, die den einrichtungsübergreifenden Austausch und das Verarbeiten von Daten transparent und nahtlos ermöglichen. Föderative Ansätze für die Bereitstellung von FDM-Services und die Speicherung von Daten wurden in Baden-Württemberg bereits erprobt und angewandt. Beispiele für etablierte Projekte sind insbesondere die Infrastruktur im Bereich High Performance Computing (z. B. *bwHPC*) und datenintensiver Dienste (z. B. *bwSFS*, *LSDF*) sowie cloudbasierter Infrastruktur (z. B. *bwCloud*). Auf Basis der gewonnenen Erfahrungen sollen strategisch die Weiterentwicklung und der Ausbau von föderativen Ansätzen von Forschungsdatenmanagement vorangetrieben und unter Berücksichtigung von bevorstehenden neuen Technologien von Grund auf ertüchtigt werden.

Entscheidende Voraussetzung für das Gelingen kooperativer Ansätze ist die vertrauensvolle Zusammenarbeit zwischen den beteiligten Einrichtungen. Gefördert wird diese unter anderem auch durch die gemeinsame Arbeit der Hochschulen in der BaWü-Datenföderation. In dieser wird kooperativ die technische Basis etabliert, um IT-Services für den gesamten Forschungsdatenlebenszyklus bereitstellen zu können.²⁶ In diesem Zusammenhang ist das Begleitprojekt *bwHPC-S5* von zentraler Bedeutung, welches die Einrichtung von Kompetenzzentren mit besonderer Expertise fördert und Beratung und Schulungen zur Nutzung der Systeme anbietet. Jedoch können die steigenden und sich verändernden Bedürfnisse von immer mehr Wissenschaftsdisziplinen über die bestehenden Angebote (z. B. Begleitprojekt *bwHPC-S5*) allein nicht adressiert werden, da die Einbindung von neuen, fachdisziplinären Anforderungen in die bestehenden Systeme besondere Expertise erfordert. In den HPC-Kompetenzzentren kümmern sich bereits jetzt sogenannte *Tiger-Teams* um die Entwicklung und Implementierung von Lösungen für besondere Einzelanforderungen von Forschungsseite. Strategisch sollte ein Ausbau der *Tiger-Teams* auch für fachspezifische Fragen des Forschungsdatenmanagements innerhalb der BaWü-Datenföderation in Erwägung gezogen werden. Damit Forschende die IT-Infrastruktur optimal nutzen (können), ist der Einsatz von hierfür geschulten „Data Stewards“ die einen entsprechenden fachlichen Hintergrund aufweisen, unbedingt erforderlich. Ebenso sind die vielfältigen Möglichkeiten der Anwendung von künstlicher Intelligenz bzw. maschinellem Lernen für die

²⁶Die technische Basis umfasst Infrastruktur für die Bereiche High Performance Computing (HPC), Data Intensive Computing (DIC) und Large Scale Scientific Data Management (LS²DM) sowie Archivstrukturen (z. B. *bwDataArchive*).

Prozessierung von Forschungsdaten ein zunehmend relevanter Aspekt; hier ist beispielsweise die automatisierte Annotation von Metadaten zu nennen.

Für die erfolgreiche Umsetzung von föderativen Ansätzen ist eine stabile, effiziente und leistungsfähige Datennetzinfrastruktur eine wichtige und notwendige Voraussetzung. Das Land verfügt hierfür mit der landeseigenen Datennetzinfrastruktur des Wissenschaftsnetzes *BelWü* über hervorragende Ausgangsbedingungen. Im Hinblick auf Informationssicherheit und zur flexiblen Definition von virtuellen Datennetzbereichen (z. B. Software Defined Networking), die datennetztechnisch die Bedürfnisse von Forschungsverbänden abbilden, müssen diese jedoch weiterentwickelt werden.

Zahlreiche Herausforderungen rund um das Forschungsdatenmanagement sind zudem weiterhin auf der lokalen Ebene der Hochschulen zu lösen, bspw. die Entwicklung von Datenpolicies. Zu berücksichtigen ist dabei insgesamt, dass in den einzelnen Einrichtungen signifikante Unterschiede bei den technischen Möglichkeiten und entsprechend auch beim Umgang mit Forschungsdaten bestehen. Um sich im internationalen Wettbewerb weiter behaupten zu können, sind Ressourcen sowohl einrichtungsübergreifend als auch lokal, sowohl für Infrastruktur als auch Personal vorzusehen.

Finanzierung der Grundinfrastruktur

Die IT-Infrastruktur und -Dienste der Hochschulen in Baden-Württemberg werden nicht nur von den eigenen Wissenschaftlerinnen und Wissenschaftlern eingesetzt: Gemeinsame Forschungsprojekte verschiedener baden-württembergischer Akteure sowie Vorhaben, deren Partner deutschlandweit oder international verteilt sind, benötigen entsprechende Ressourcen für Forschungsdatenmanagement. Dafür können in der Regel keine eigenen IT-Infrastrukturen und -Dienste projektbezogen aufgebaut werden.

Insbesondere für drittmittelfinanzierte Projekte gilt dabei, dass Sachmittel für die Nutzung von technischer Infrastruktur in der Regel vonseiten der Mittelgeber nicht pauschal gefördert werden (dürfen), da die Bereitstellung von zeitgemäßen IT-Infrastrukturen und -Diensten zu den Grundaufgaben der Forschungseinrichtungen zählt.²⁷ Die für die Übertragung und Verarbeitung auch von umfangreichen Datensätzen und vor allem auch die lokale, langfristige Speicherung von Forschungsdaten nötigen Kapazitäten müssen daher in der Regel durch die einzelnen Einrichtungen grundfinanziert werden. Dies gilt auch für die Arbeit der Konsortien der NFDI, welche in erster Linie personelle Ressourcen fördert und deren Tools auf der vorhandenen IT-Infrastruktur und -Diensten aufbauen sollen. Umgekehrt bedeutet dies: Sollten die Einrichtungen nicht in der Lage sein, eine den aktuellen Erfordernissen genügende technische Infrastruktur vorzuhalten, könnte dies zu einem reduzierten Drittmittelaufkommen führen.

Dem forschungsstarken Land Baden-Württemberg mit seiner reichen Hochschullandschaft fällt daher die verantwortungsvolle Rolle zu, die für ein nachhaltiges Forschungsdatenmanagement notwendigen IT-Infrastrukturen und -Dienste in der benötigten Kapazität und Qualität bereitzustellen und zu finanzieren. Dabei gilt es nicht nur, die initialen

²⁷Vgl. etwa die Förderrichtlinien der Deutschen Forschungsgemeinschaft (DFG): <https://www.dfg.de/de/grundlagen-themen/grundlagen-und-prinzipien-der-foerderung/forschungsdaten/beantragbare-mittel> (29.09.2023).

Anschaffungskosten zu berücksichtigen, sondern auch die nötigen Investitionen in Instandhaltung und Aktualisierung von Systemen einzukalkulieren. Für ein nachhaltiges Forschungsdatenmanagement bedarf es daher langfristig erheblicher Investitionen sowohl in Hard- und Software als auch qualifiziertes Personal.

Themenfelder

Speicherung der Daten: Hot und Cold Data

Die technische Umsetzung der Speicherung von Forschungsdaten unterscheidet *Hot Data* von *Cold Data*: Als *Hot Data* bezeichnet man gespeicherte Datensätze, auf die schnell und regelmäßig zugegriffen werden muss, weil beispielsweise aktuell mit ihnen gearbeitet wird. Unter *Cold Data* werden bereits verarbeitete und archivierte Daten verstanden, die seltener für Forschungsaktivitäten herangezogen werden.

Die Speicherdienste für *Hot Data* sollen die Verfügbarkeit der Daten in Echtzeit für einzelne Forschende, aber auch für kooperative Forschungsverbände gewährleisten. Großforschungsgeräte wie z. B. Hochleistungsrechner, High-Throughput-Mikroskope, Hardware für hochauflösendes Videomaterial oder Sequencer, generieren standortabhängig erhebliche Mengen an *Hot Data*. Diese müssen entsprechend an Datenspeicher übertragen werden. In diesem Kontext kann das Datennetz zum Flaschenhals werden. Deshalb sind dezentrale Pufferspeicher erforderlich, damit große Datenmengen zeitversetzt in die Datenspeicher eingespeist werden können. Für die Bereitstellung der Daten sind Repositorien zu etablieren. In den Bereichen, in denen erhebliche Mengen an Daten gespeichert werden, ist die Verwendung semantischer bzw. KI-basierter Webansätze, die das Durchforsten und das Abrufen der benötigten Daten ermöglichen, eine Grundvoraussetzung für die tägliche Arbeit der Forschenden.

Repositorien können auch der zentrale Speicherort für *Cold Data* sein. Besonders wichtig für die Auffindbarkeit sind persistente Identifikatoren und Metadaten, die zwar community-abhängig, aber standardisiert sind. Repositorien für *Cold Data* verwenden weniger performante Speichersysteme, um Datenanfragen zu bearbeiten. Dadurch und aufgrund der nicht notwendigen Echtzeitverfügbarkeit fallen die Gesamtkosten bezogen auf die Kapazität im Vergleich zu *Hot Data* deutlich geringer aus. Sie sind aber aufgrund der insgesamt immer weiter steigenden Menge der vorhandenen Daten nicht zu unterschätzen.

Datenzugriffe: Rechte und Rollenkonzepte

Mit dem Übergang von *Hot* zu *Cold Data* rücken langfristige Speicherung und technische Nachnutzbarkeit in den Fokus. Hier sind Entscheidungen in Bezug auf eine angepasste Speicherlösung und ein entsprechendes Rechte- und Rollenkonzept für die Nachnutzung entscheidend. Auch wenn die hierfür notwendigen Konzepte bereits teilweise umgesetzt worden sind (etwa im Rahmen von *bwIDM*), ist ein Ausbau – mit Berücksichtigung auch von bislang noch nicht im Fokus befindlichen Fachdisziplinen – sowohl strategisch als auch hinsichtlich der materiellen und personellen Ressourcen erforderlich.

Langzeitarchivierung von Forschungsdaten

Durch die Vorgaben von Forschungsförderern und Forschungseinrichtungen werden Forschende aller Fachrichtungen verpflichtet, ihre Forschungsdaten nach den FAIR-Prinzipien für mindestens zehn Jahre zu archivieren. Das Forschungsdatenmanagement und insbesondere die Langzeitarchivierung von Forschungsdaten stellen grundsätzlich andere Anforderungen an die benötigte IT-Infrastruktur als etwa das Hoch- und Höchstleistungsrechnen. Denn ein Ausfall der FDM-IT-Infrastruktur kann zu einem irreversiblen Verlust wichtiger Daten führen, wohingegen im HPC-Umfeld entsprechende Berechnungen notfalls wiederholt werden können. Für die Langzeitarchivierung von Forschungsdaten sind daher spezifische Technologien vorzusehen, die robust und redundant angelegt sind.

Ein landeseinheitliches Konzept für die Langzeitarchivierung der unterschiedlichen Arten von Forschungsdaten ist eine Notwendigkeit. Hierbei sollte sorgfältig geprüft werden, welche Forschungsdaten für wie lange und mit welcher Sicherheitsstufe gespeichert werden müssen. Während ein Großteil der Daten nach einer gewissen Zeit gelöscht werden kann, gibt es auch Daten, die auf Dauer erhalten werden müssen. Bei der Archivierung geht es nicht nur um die „reinen“ Daten (sog. Bitstream-Preservation), sondern auch um die nötigen Metadaten und Kontexte wie beispielsweise Forschungssoftware, um die Daten langfristig nutzbar zu halten. Die Etablierung von kooperativen Infrastrukturansätzen auf Landesebene ist im Zusammenhang mit der Langzeitarchivierung konzeptionell noch zu definieren und entsprechend umzusetzen.

Software für Forschungsdatenmanagement und Virtuelle Forschungsumgebungen

Gutes Forschungsdatenmanagement ist ohne die passende Softwareunterstützung nicht denkbar. Der Bereitstellung von entsprechenden Tools, z. B. für die Visualisierung von Daten, kommt daher große Bedeutung zu. Dabei ist zwischen generischen Lösungen, z. B. zur Planung und Verwaltung des Forschungsprojekts, und fachspezifischer Unterstützung, z. B. zu Elektronischen Laborbüchern (Electronic Lab Notebooks – ELN), zu unterscheiden. Auch die Lizenzierung spielt für den Einsatz von FDM-Software eine Rolle, da es in diesem Bereich sowohl Open-Source-Entwicklungen als auch nicht-freie Software gibt.

Die Nutzung und Entwicklung von spezifischer Software für Forschungsdatenmanagement spielt im Prozess der Datenerhebung und der Datenanalyse eine zentrale Rolle. Je nach Disziplin und Fragestellung erfordert die Datenanalyse ein spezifisches Know-How, das sich in zusätzlichem Bedarf an Ressourcen aus dem Bereich des wissenschaftlichen Rechnens widerspiegelt. Nicht selten werden in Forschungsvorhaben sensible personenbezogene Daten verarbeitet. Hierfür sind weitere Anforderungen im Hinblick auf Anonymisierung oder Pseudonymisierung zu berücksichtigen. Diese erfordern auch spezifische technische Maßnahmen, um IT-Compliance und die teilweise notwendigen Zertifizierungen für Forschungsaktivitäten zu erreichen. Welche Zertifizierungen für Forschungsaktivitäten sinnvoll sind, muss dabei auch mit Blick auf die Wettbewerbsfähigkeit des Landes berücksichtigt werden.

Um den aktuellen Anforderungen aus der (internationalen) Forschungsförderung sowie den Policies der wissenschaftlichen Einrichtungen gerecht zu werden, können Forschende teilweise schon jetzt auf an ihre Bedürfnisse angepasste Werkzeuge zugreifen. Bei der Erstellung von fachspezifischen Datenmanagementplänen unterstützen bspw. Softwarelösungen wie RDMO²⁸. Ungeachtet der stetigen Optimierung der Unterstützungstools ist die Beratung und Begleitung der Forschenden durch wissenschaftlich und technisch qualifizierte Expertinnen und Experten im gesamten Forschungsdatenlebenszyklus für das erfolgreiche Forschungsdatenmanagement unverzichtbar.

Der Aufbau und Ausbau virtueller Forschungsumgebungen (VFU) wird eine der zentralen Aufgaben für die wissenschaftlichen Infrastruktureinrichtungen der kommenden Jahre sein. VFUs weisen eine große fachspezifische Varianz auf, sie eignen sich für die wissenschaftliche Arbeit am Forschungsobjekt selbst (z. B. Aufarbeitung von Genom-Daten, Annotieren eines Linguistik-Korpus), eine besondere Bedeutung kommt ihnen aber durch die Möglichkeiten des kollaborativen wissenschaftlichen Arbeitens zu.

Forschungssoftware

Um Forschungsergebnisse reproduzierbar zu machen und zu halten, muss auch der Zugang zu Forschungssoftware dauerhaft gewährleistet sein. Verwendete Software sollte grundsätzlich kenntlich gemacht und der Quellcode von im Forschungsprozess entwickelter Software nach den Regeln der guten wissenschaftlichen Praxis festgehalten werden. Dies gilt insbesondere bei der Archivierung von Software, im Zuge derer auf geeignete Formate und eine ausreichende Dokumentation geachtet werden muss. Die technischen Möglichkeiten, um Forschungssoftware angemessen zu archivieren, sind grundsätzlich vorhanden. Insbesondere Fragen des rechtlich-organisatorischen Rahmens sind jedoch weiterhin offen: So liegen für quelloffene Software bereits publizierte Empfehlungen zur Vergabe geeigneter Lizenzen und Anforderungen an die Versionskontrolle und Dokumentation vor. Ebenso haben verschiedene Einrichtungen Policies zum Umgang mit sowie zur Pflege von Softwarepaketen erlassen; auf bestehenden Best Practices aufbauende Standards für den nachhaltigen Umgang mit Forschungssoftware sind jedoch noch nicht flächendeckend eingeführt.

Informationssicherheit und Schutz der Forschungsdaten

Informationssicherheit ist eine zentrale Herausforderung für das Forschungsdatenmanagement, sowohl für *Hot* wie auch für *Cold Data*. Deshalb sind Vorkehrungen zu treffen, um den Schatz an Forschungsdaten im Land Baden-Württemberg bestmöglich zu schützen. Auch hierfür sind bereits erste Strukturen aufgebaut worden (z. B. *bwInfoSec*), die aber mit den wachsenden Herausforderungen nur eingeschränkt Schritt halten können.

Ein besonders hohes Schutzniveau ist bei der Verarbeitung sensibler Daten (insbesondere aus Bereichen wie Medizin oder Psychologie) zu beachten. Dies bringt zusätzliche technische Herausforderungen mit sich, die im Rahmenkonzept HPC / DIC adressiert werden.²⁹ Ein

²⁸<https://rdmorganiser.github.io> (29.09.2023).

²⁹Rahmenkonzept der Universitäten des Landes Baden-Württemberg für das High-Performance Computing (HPC) und Data-Intensive Computing (DIC) für den Zeitraum 2025 bis 2032.

einerseits striktes, andererseits auch flexibles Rollen- und Rechtemanagement (z. B. über *bwIDM*) für den Zugang zu den Daten ist dazu notwendig. Es ist wichtig zu beachten, dass die wissenschaftlichen Aktivitäten nicht nur im Landeskontext, sondern auch im Rahmen von nationalen und internationalen Kooperationen erfolgen.

Open Data und Nachhaltigkeit

Ein nachhaltiges Forschungsdatenmanagement sichert die möglichst offene Verfügbarkeit der Daten. Open Data bedeutet dabei keineswegs eine grenzenlose Verfügbarkeit aller Daten. Sensible Daten, die z. B. aus Gründen des Datenschutzes besonders schutzwürdig sind, müssen auch im Rahmen einer Open-Data-Strategie angemessen gesichert sein. Die Architektur der IT-Infrastruktur und -Dienste muss diese Anforderungen erfüllen. Manche Forschungscommunities verlangen Zertifizierungen für zugrundeliegende Infrastruktur, ohne welche das Speichern und die Verarbeitung von Forschungsdaten nicht zulässig ist. Auf Landesebene ist entsprechend darauf zu achten, dass diese Rahmenbedingungen für die IT-Infrastruktur berücksichtigt werden, damit die Einrichtungen des Landes nicht aus diesen Forschungsaktivitäten ausgeschlossen werden. Gleiches gilt auch für die Bereitstellung von FDM-Software sowie Forschungssoftware.

Zudem wächst in Anbetracht von Klima- und Energiekrise die Verantwortung für ökologische Nachhaltigkeit. Mit der Green-IT Strategie³⁰ nimmt die Landesregierung bewusst auch die Rechenzentren und dadurch die IT-Infrastruktur für Forschungsdaten in den Blick. Durch die Umsetzung von durchdachten Lösstrategien können Speicherkapazitäten ressourcenschonender verwaltet und eingespart werden. Langfristig sollte für die Betreiber-Einrichtungen unbedingt eine Zertifizierung zur Energie-Effizienz angestrebt werden.

Empfehlungen

- Der Aufbau weiterer lokaler oder gegebenenfalls regionaler Forschungsdatenrepositorien in Ergänzung zu den neuen, übergreifenden Strukturen wie der NFDI und anderer fachlicher Initiativen ist empfehlenswert. Dies gilt auch für Speicherstrukturen für Forschungssoftware.
- Zur Verarbeitung, Speicherung und Langzeitarchivierung von Forschungsdaten sind Investitionen in hoch verfügbare Infrastrukturen notwendig, die bereits an größeren Einrichtungen etabliert sind (z. B. *LSDF*, *bwSFS*). Dies soll auch über das bestehende, von der DFG bereits positiv begutachtete Rahmenkonzept *bwDATA* erfolgen.
- Ein landeseinheitliches Konzept für die Langzeitarchivierung der unterschiedlichen Arten von Forschungsdaten und Forschungssoftware sollte entwickelt werden. Die Einrichtung eines verteilten Landesdatenarchivierungszentrums sollte geprüft werden.

³⁰<https://um.baden-wuerttemberg.de/de/ministerium/aufgaben-organisation/nachhaltige-landesverwaltung/green-it> (29.09.2023) bzw. https://um.baden-wuerttemberg.de/fileadmin/redaktion/m-um/intern/Dateien/Dokumente/1_Ministerium/Aufgaben_und_Organisation/Green-IT/191122-Kurzbericht-Landesstrategie-Green-IT.pdf (13.10.2025).

- Die lokale Anschlussfähigkeit sowie die Nutzbarkeit dieser regionalen Infrastrukturen ist sicherzustellen. Für die Hochschulen für angewandte Wissenschaften und andere kleinere und mittelgroße Einrichtungen wird die Nutzungsmöglichkeit dieser landeseigenen Infrastrukturen empfohlen.
- Die kooperative regionale BaWü-Datenföderation ist weiter zu etablieren und als Basis für den gesamten Forschungsdatenlebenszyklus auszubauen.
- Der Ausbau von sogenannten *Tiger-Teams* für FDM-spezifische Fragen innerhalb der BaWü-Datenföderation sollte in Erwägung gezogen und der Einsatz von „Data Stewards“ gefördert werden.
- Speziell für den Umgang mit hoch sensiblen Daten ist das Konzept der BaWü-Datenföderation anzupassen, etwa über organisatorische (wie Datentreuhandstellen) und technische Lösungen (geeignete technische Verfahren zur Anonymisierung, Pseudonymisierung etc.).