



Leitfaden – Verantwortungsvoller Umgang mit Forschungsdaten

Erarbeitet vom Arbeitskreis Forschungsdatenmanagement (AK FDM)

Charakter des Papiers: Best-Practice-Empfehlungen differenziert nach Stakeholdern

Adressaten: Forschende, Infrastruktureinrichtungen, Hochschulen

Autorinnen und Autoren: Dirk von Suchodoletz, Peter Brettschneider, Elisabeth Böker, Jochen Apel, Dorothee Iglezakis, Karsten Schmidt, Gabriel Schneider

Inhalt

Leitfaden – Verantwortungsvoller Umgang mit Forschungsdaten	1
Einleitung	2
1. Rahmenbedingungen.....	2
1.1 Forschungsdaten-Policies	3
1.2 Gute wissenschaftliche Praxis und Ethik.....	3
1.3 Rechtliche Rahmenbedingungen.....	5
1.4 IT-Services und Sicherheit	6
2. Stakeholder	7
2.1 Definition der Stakeholder	7
2.2 Rollen / Aufgaben / spezifische Herausforderungen	7
3. Umsetzung.....	9
3.1 Kernentscheidungen einer Forschungsdaten-Governance	9
3.2 Qualitätsmanagement.....	9
3.3 Daten-Kuratierung.....	10
3.4 Rechtemanagement	11
4. Kosten	11
Ausblick	13

Einleitung

Forschungsdaten sind Ausgangs- und langfristig aufzubewahrendes Endprodukt wissenschaftlicher Forschung. Die Vielzahl der in den Forschungsprozess involvierten Beteiligten und die für den Wissenschaftsbetrieb typische, hohe Personalfuktuation machen das Management dieser Daten zu einem komplexen Prozess, der mit erheblichen Herausforderungen einhergeht.

Die Notwendigkeit, digitale Forschungsprimärdaten über lange Zeiträume zu erhalten, ist dabei unbestritten. Ein klares Zeichen dafür wurde – angestoßen durch die Gemeinsame Wissenschaftskonferenz von Bund und Ländern (GWK) – mit der Planung und Umsetzung einer Nationalen Forschungsdateninfrastrukturen (NFDI) gesetzt. Diese schafft wichtige Impulse für den wissenschaftlichen Fortschritt und neue Innovationen, die sich aus der gemeinsamen interdisziplinär entstehenden Struktur ergeben können.

Dabei bezieht der Prozess des Erhalts eine Vielzahl von Akteuren (Hochschule, untergeordnete Einrichtungen, Forschende, Infrastruktureinrichtungen, Dritte) ein, die bei der Planung, Erfassung, Dokumentation und Archivierung beteiligt sind. Diese zumeist über unterschiedliche Strukturen mit vielfältigen Verpflichtungen realisierte Aufgabe benötigt verantwortungsvolle Begleitprozesse der eigentlichen Forschungstätigkeit. Die Bedeutung, schon in der Planung Kernaspekte eines nachvollziehbaren Managements der Forschungsdaten zu etablieren, wird sehr deutlich, wenn einer oder mehrere beteiligte Akteure für die entsprechende Dokumentation des Forschungsprozesses bei der Erzeugung und Verarbeitung des Forschungsdatums nicht mehr zur Verfügung stehen. Die somit entstehende Lücke in der Nachvollziehbarkeit der Forschungsarbeit führt zumeist zu Forschungsdaten, die für sich keine Eigenständigkeit besitzen, was ihre Archivwürdigkeit im Sinne der FAIR-Prinzipien einschränkt.

An vielen Verfahrensketten ist der Mangel an Struktur sehr deutlich: So wird in überraschend vielen Projektanträgen eine Dokumentation der Forschungsdaten erst gegen Ablauf der Förderperiode angedacht. Der subjektive Mehraufwand wird zusätzlich gesteigert durch die ungenaue Erinnerung bzw. durch die fehlenden Akteure, die wesentliche Komponenten der Dokumentation nicht mehr bzw. nur sehr schwer reproduzierbar machen. Eine strukturierte Einbindung aller Akteure über den gesamten Lebenszyklus der Daten hinweg erleichtert nicht nur die kontrollierte Wissensspeicherung in der eigenen Arbeitsgruppe, sondern auch die Kommunikation der Forschungsdaten in der Fach-Community. Die Verantwortung für Daten (und Software) muss daher in geregelter Form zu jedem Zeitpunkt im Datenlebenszyklus durch geeignete Akteure ausgeübt werden. Dieses Dokument stellt eine Handreichung für praktische Aspekte des Forschungsdatenmanagements [FDM] sowohl für einzelne Forschende und Arbeitsgruppen bis hin zu größeren Verbundprojekten dar. Ziel ist es, eine Grundlage zu bieten, um den verschiedenen Stakeholdern Orientierung im Umgang mit Forschungsdaten zu geben und Best-Practices für alle Phasen im Datenlebenszyklus vorzuschlagen. Dazu werden Rahmenbedingungen [1.], Stakeholder [2.], Umsetzung [3.] und Kosten [4.] eines verantwortungsvollen Umgangs mit Forschungsdaten beleuchtet und konkrete Empfehlungen ausgesprochen.

1. Rahmenbedingungen

Mit der zunehmenden Verbreitung einer guten wissenschaftlichen Praxis im Zusammenhang mit digitalen Forschungsdaten ist der Aufbau und die Fortentwicklung eines institutionalisierten FDM

nicht mehr optionaler Zusatzaufwand, sondern Kernaufgabe von Wissenschaftseinrichtungen. Zunehmend erwarten Forschungsförderer auf Projektebene Angaben zum FDM als Voraussetzung für die Vergabe von Zuwendungen. Parallel dazu besteht in vielen Disziplinen das Bestreben, gut kuratierte und annotierte Datenpublikationen den traditionellen Publikationen gleichzustellen.

1.1 Forschungsdaten-Policies

Empfehlung	Adressat
Institutionelle Forschungsdaten-Policy erlassen bzw. weiterentwickeln.	Hochschulen
Gestaltung der Forschungsdaten-Policy in den Gremien der akademischen Selbstverwaltung. Umsetzung dieser Leitlinien in der Forschungspraxis.	Forschende
Unterstützung bei Ausarbeitung und Umsetzung der Forschungsdaten-Policy.	Infrastruktureinrichtungen

Der verantwortungsvolle Umgang mit Forschungsdaten erfordert eine Abstimmung mit allen beteiligten Akteuren, wie Hochschulen, untergeordnete Einrichtungen, Forschende, Infrastruktureinrichtungen und Dritte. Eine gute Orientierung bietet die Vorgehensweise von RISE-DE, die dieses bereits erfolgreich umgesetzt hat.¹ Am Anfang steht die Erstellung und Verabschiedung einer Leitlinie oder Policy in den Gremien der akademischen Selbstverwaltung auf der Ebene der Gesamtorganisation. Je nach Ausgestaltung und konkreten Anforderungen sollten Fakultäten und Fachbereiche diesen allgemeinen Rahmen in zusätzlichen, fachspezifischen Leitlinien konkretisieren.

Eine institutionelle FDM-Policy umfasst grundlegende Festlegungen und Regelungen für das FDM an einer Hochschule und ihren nachgeordneten Entitäten. Unter anderem sollten die verantwortlichen Akteure und ihre Rollen benannt, eine Leitlinie innerhalb der Hochschule entworfen und Vorgaben zu Dauer und Modalitäten der Aufbewahrung von Forschungsdaten enthalten sein. Dieses sollte darauf abzielen, die Idee von Open Science zu fördern, den Kriterien der FAIR-Prinzipien zu folgen und potentielle Konfliktfälle von vorneherein zu entschärfen. Zudem sollte dieses Vorgehen die Interessen der verschiedenen Beteiligten von vorne herein zum Ausgleich bringen. So könnte z.B. eine bestimmte Art der Vergabe von Nutzungsrechten und Lizenzierung empfohlen werden.

1.2 Gute wissenschaftliche Praxis und Ethik

Empfehlung	Adressat
Gute wissenschaftliche Praxis und ethische Forschung befördern und institutionell absichern.	Hochschulen

¹ Vgl. Niklas K. Hartmann u.a.: RISE-DE – Referenzmodell für Strategieprozesse im institutionellen Forschungsdatenmanagement, <https://doi.org/10.5281/zenodo.2549344>

Regeln der guten wissenschaftlichen Praxis und Wissenschaftsethik beachten.	Forschende
Ggf. Beratung anbieten, speziell auch zu den Aspekten rund um Forschungsdaten und bei Konfliktfällen vermitteln.	Infrastruktureinrichtungen

Die DFG-Leitlinien zur Sicherung guter wissenschaftlicher Praxis² setzen entscheidende Standards sowohl für den Inhalt von FDM-Policies als auch für den verantwortungsvollen Umgang mit Forschungsdaten an sich. Hochschulen und Forschende verpflichten sich auf diese z.B. bei Annahme von Fördermitteln der DFG. Nicht weniger bedeutsam ist die Beachtung wissenschaftsethischer Gebote.

So verpflichten etwa die Leitlinien zur Sicherung guter wissenschaftlicher Praxis der DFG die Wissenschaft u.a. zu einer nachvollziehbaren, den Standards der jeweiligen Fach-Community entsprechenden Dokumentation.³ Auch auf die FAIR-Prinzipien nehmen die Leitlinien explizit Bezug und fordern – sofern keine gewichtigen Gründe entgegenstehen – dazu auf, Forschungsdaten zugänglich zu machen.⁴ Nicht zuletzt wird die Bedeutung der langfristigen Archivierung von Forschungsdaten, -ergebnissen und -software betont.⁵

Hauptadressaten der Gebote der guten wissenschaftlichen Praxis und wissenschaftsethischer Regeln sind die Forschenden selbst. Infrastruktureinrichtungen kann im Hinblick auf Forschungsdaten ggf. eine beratende und moderierende Rolle zukommen. Demgegenüber obliegt es Hochschulen Rahmenbedingungen zu schaffen, die gute wissenschaftliche Praxis befördern und institutionell absichern (z.B. durch Ethikkommissionen).

Insofern besteht eine Wechselwirkung zwischen den Standards guter wissenschaftlicher Praxis und universitären Forschungsdaten-Policies. Letztere referenzieren die gute wissenschaftliche Praxis nicht nur, sondern konkretisieren idealerweise deren Anforderungen auf einer lokalen Ebene.

² Vgl. DFG: Leitlinien zur Sicherung guter wissenschaftlicher Praxis, 2019, https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf [9.11.2020].

³ Vgl. DFG: Leitlinien zur Sicherung der guten wissenschaftlichen Praxis, 2019, Leitlinie 12, https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf. Ausführlicher dazu: <https://www.forschungsdaten.info/themen/ethik-und-gute-wissenschaftliche-praxis/gute-wissenschaftliche-praxis-und-fdm/> [9.11.2020].

⁴ Vgl. DFG: Leitlinien zur Sicherung der guten wissenschaftlichen Praxis, 2019, Leitlinie 13, https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf [9.11.2020].

⁵ Vgl. DFG: Leitlinien zur Sicherung der guten wissenschaftlichen Praxis, 2019, Leitlinie 17, https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf [9.11.2020].

1.3 Rechtliche Rahmenbedingungen

Empfehlung	Adressat
Verbindliche Vereinbarungen auf Projekt- und Lehrstuhlebene befördern. FDM-Probleme in standardisierten Arbeitsverträgen und Promotionsvereinbarungen berücksichtigen.	Hochschulen
Frühzeitige Berücksichtigung rechtlicher Fragen.	Forschende
Beratung und Schulungen anbieten – ggf. in Kooperation mit anderen Stellen wie der Rechtsabteilung.	Infrastruktureinrichtungen

Der verantwortungsvolle Umgang mit Forschungsdaten wird zum einen vom für Daten geltenden Rechtsregime und zum anderen durch die auf Forschende anwendbaren rechtlichen Regelungen beeinflusst.

Ersteres umfasst insbesondere urheber- und datenschutzrechtliche Regelungen. Insofern besteht ein Spannungsverhältnis zur wissenschaftspolitischen Zielsetzung einer möglichst großen Transparenz und Offenheit von Forschung.⁶ Seine Auflösung erfordert einen vorausschauenden Umgang mit den Rechtspositionen an Daten und besondere Umsicht im Umgang mit sensiblen Daten. Regelmäßig gestellte rechtliche Fragen sind u.a.:

- Wem „gehören“ Forschungsdaten? Wer darf darüber entscheiden, ob Daten publiziert werden bzw. wer Zugang hat?
- Welchen rechtlichen Einschränkungen unterliegt die Nutzung übernommener Forschungsdaten?
- Unter welchen rechtlichen Bedingungen dürfen bzw. sollen Daten öffentlich gemacht werden?
- Was gilt für die Verarbeitung personenbezogener Daten? Dürfen solche Daten archiviert oder publiziert werden?

Die mit diesen Fragen verbundenen Problemstellungen und Konflikte lassen sich deutlich leichter entschärfen, wenn sie bereits in einem frühen Stadium der Planung bedacht und klare rechtliche Regelungen zwischen den Beteiligten getroffen werden. Nicht zuletzt muss aber das Bewusstsein für rechtliche Fragen durch Schulungen und Beratungen geschärft und ihre Einhaltung eingefordert werden. Dabei sind nicht nur Infrastruktureinrichtungen und andere Stellen der Universität gefordert entsprechende Angebote zur Verfügung zu stellen, sondern auch die Forschenden, die aktiv Beratung suchen und für rechtliche Themen offen sein sollten.

⁶ Vgl. Wissenschaftsrat: Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020 https://www.wissenschaftsrat.de/download/archiv/2359-12.pdf?__blob=publicationFile&v=1 [9.11.2020]; siehe auch: Open Access zu Forschungsdaten: <https://idw-online.de/de/news490437> bzw. RfII „Stellungnahme des Rates für Informationsinfrastruktureinrichtungen (RfII) zu den aktuellen Entwicklungen rund um Open Data und Open Access“, März 2019. <http://www.rfii.de/?p=3748> [9.11.2020].

Zweitens sind die auf Forschende anwendbaren Regelungen jedoch ein unverzichtbarer Teil des verantwortungsvollen Umgangs mit Forschungsdaten. So werden konkrete FDM-Verpflichtungen in Förderverträgen festgeschrieben. Weiterhin können arbeits- und IT-Sicherheitsrechtliche Belange eine Rolle spielen. Konfliktpotentiale sollten von Beginn an bedacht und durch dokumentierte Entscheidungen (z. B. vertragliche Regelungen) entschärft werden. Beispielsweise bietet es sich an in Arbeitsverträgen oder Promotionsvereinbarungen zu regeln, ob Forschende, die ein Projekt verlassen, weiterhin Zugriff auf die Daten haben dürfen, um z. B. Publikationen fertigstellen zu können. Dabei können technischer Datenschutz und berechnigte Forschungsinteressen durchaus in Widerspruch stehen.

1.4 IT-Services und Sicherheit

Empfehlung	Adressat
Bereitstellung eines Portfolios an grundlegenden, sicheren FDM-Services	Hochschulen
Services mit hoher Sicherheit verwenden	Forschende
Aufbau und Betrieb von FDM-Diensten. Dabei sollten zugleich Strukturen zur Qualitätssicherung geschaffen werden.	Infrastruktur

FDM- und Speichersysteme sind an vielen Universitäten bei der Universitätsbibliothek beziehungsweise dem Rechenzentrum lokalisiert; es können aber auch weitere Einrichtungen involviert sein. Zudem haben sich in manchen Fach-Communities Institutionen und Dienste etabliert, die u. a. international anerkannte Repositorien betreiben.⁷ Befördert durch die NFDI ist hier mit weiteren Entwicklungen zu rechnen.

Eine Hochschule sollte als Organisation ein Portfolio an grundlegenden FDM-Diensten anbieten, um im Zeitalter der Digitalisierung als relevanter Akteur wahrgenommen zu werden. Auf der institutionellen FDM-Policy sollte eine Dienststruktur aufgebaut werden, die festlegt, nach welchen Workflows und auf welchen Systemen Daten verarbeitet werden, damit alle Phasen im Lebenszyklus von Daten durch die Forschung konkret ausführbar sind (vgl. Abbildung 1).

Systeme für das FDM sollten zudem technische Strukturen zur Kuratierung und Qualitätssicherung von Datensätzen bereitstellen. Insbesondere ist für institutionelle Repositorien die Etablierung eines Peer-Review-Verfahrens für Forschungsdaten zu empfehlen, dass nach formalen und bzw. oder inhaltlichen Kriterien die Aufnahme von Forschungsdaten begleitet.

Umgekehrt obliegt es Forschenden, bei der Verarbeitung ihrer Forschungsdaten IT-Sicherheit mitzudenken und solche Dienste zu meiden, bei denen diesbezüglich Bedenken angebracht sind.

⁷ Ein Beispiel ist die an der Universität Konstanz betriebene MovebankData Repository:
<https://www.datarepository.movebank.org/>, [11.11.2020].

2. Stakeholder

In den Leitlinien zur Sicherung der guten wissenschaftlichen Praxis der DFG heißt es, dass die Rollen und Verantwortlichkeiten der beteiligten Personen zu jedem Zeitpunkt des Forschungsprozesses geklärt sein müssen.⁸

2.1 Definition der Stakeholder

Dabei sind Stakeholder an der eigenen Hochschule bzw. Forschungseinrichtung ebenso zu berücksichtigen wie externe Stakeholder:

- Stakeholder an einer Hochschule: Forschende, Hochschule, Fachbereiche und Fakultäten, Lehrstühle und Forschungsprojekte, Infrastruktureinrichtungen
- Externe Stakeholder: Forschungsförderer, fachspezifische Infrastruktureinrichtungen, Fach-Communities, Kooperationspartner an anderen Einrichtungen bzw. Hochschulen.

2.2 Rollen / Aufgaben / spezifische Herausforderungen

Empfehlung	Adressat
FDM-Beauftragte etablieren	Hochschulen
DMP aufsetzen, Unterstützung dazu einholen	Forschende
Unterstützung bei DMP	Infrastruktureinrichtungen

Koordination und Zusammenarbeit zwischen den verschiedenen Stakeholdern erfordern regelmäßigen Austausch und Absprachen.⁹ Dies gilt in gleicher Weise auch für die Beteiligten an einem spezifischen Forschungsprojekt. Dazu hat RISE-DE konkrete Vorschläge erarbeitet.¹⁰ Die Regelungen sollten dabei so früh wie möglich getroffen werden und idealerweise in Form eines Datenmanagementplan [DMP] dokumentiert werden. Grundlage für die Zuordnung der Rollen können Anforderungen von Forschungsförderern aber auch institutionelle oder fachspezifische Policies sein.

Für Forschungsprojekte sollte deshalb stets ein DMP erstellt werden. Dies sollte bereits in der Projektplanungsphase geschehen. DMPs sind dabei nicht nur als Management-Tool für ein Forschungsprojekt zu verstehen, sondern auch als Konkretisierung der Vorgaben von

⁸ Vgl. DFG: Leitlinien zur Sicherung der guten wissenschaftlichen Praxis, 2019, https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf, Leitlinie 7.

⁹ Vgl. DFG: Leitlinien zur Sicherung der guten wissenschaftlichen Praxis, 2019, https://www.dfg.de/download/pdf/foerderung/rechtliche_rahmenbedingungen/gute_wissenschaftliche_praxis/kodex_gwp.pdf, Leitlinie 7.

¹⁰ Vgl. RISE-DE, 10.5281/zenodo.3585556.

Forschungsförderern und institutionellen Policies in einem konkreten Forschungsvorhaben. Daher empfiehlt es sich, auf Leitlinien und Policies der Hochschule und Fakultäten zu verweisen.

Auf Ebene der Fakultäten sollten zudem "FDM-Beauftragte" etabliert werden. Zu Ihren Aufgaben könnte auch die Entscheidung gehören, ob langzeitarchivierte Forschungsdaten nach Ablauf der ursprünglich festgelegten Speicherfrist weiter vorgehalten oder gelöscht werden sollen. Eine Orientierung können dabei die Usancen des Fachgebietes und generellen Daten- und Kosteninformationen liefern. Für nicht publizierte Daten müssen zudem Regelungen hinsichtlich der Zugänglichmachung der Daten bei entsprechenden Anfragen von Dritten getroffen werden, insbesondere für den Fall, dass die ursprünglichen Datenproduzentinnen und -produzenten nicht mehr als Ansprechpartnerinnen und Ansprechpartner zur Verfügung stehen. Auch diese Rollen könnten FDM-Beauftragte in Kooperation mit den Betreibern von Repositorien oder Archivservices einnehmen.

Allen Modellen, die Lebenszyklen von Forschungsdaten betrachten, ist die Erkenntnis gemein, dass sich Rollen beim Umgang mit diesen Daten verändern. Daraus folgt, dass eine dauerhafte Fixierung von Rechten und Pflichten nur bedingt möglich ist, sondern vielmehr Gestaltungsfreiheit für situative Regelungen bleiben muss. Zusätzliche Differenzierungen werden durch die Art, den Umfang und die Disziplin(en) eines Forschungsvorhabens bestimmt.

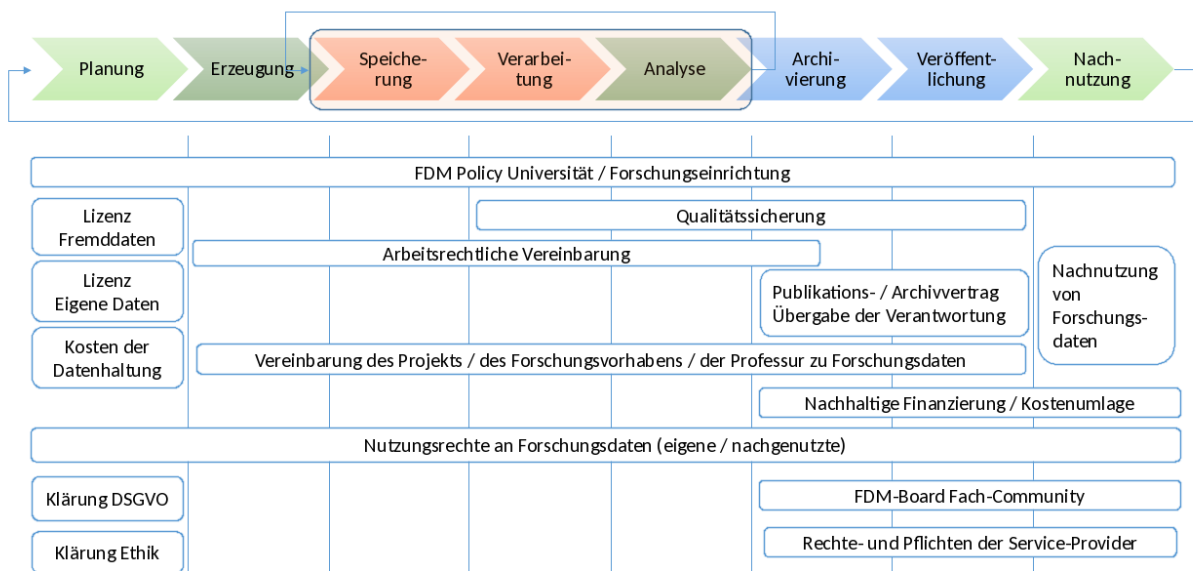


Abbildung 1: Auftretende Aspekte im Zeitablauf des Forschungsdatenlebenszyklusses

3. Umsetzung

3.1 Kernentscheidungen einer Forschungsdaten-Governance

Empfehlung	Adressat
Durchführung eines Dialogs zur Festlegung der Rahmenbedingungen von FDM und Fixierung der Ergebnisse in verbindlicher Form.	Hochschulen
Aktive Mitgestaltung der Forschungsdaten-Governance.	Forschende
Aktive Mitgestaltung der Forschungsdaten-Governance.	Infrastruktureinrichtungen

Damit Festlegungen zum Umgang mit Daten nicht als einseitig empfunden werden, müssen die zentralen Entscheidungen beim verantwortungsvollen Umgang mit Forschungsdaten auf einem möglichst breit angelegten Konsens basieren. Dabei sollten sowohl die Forschenden als auch die forschenden Gliederungen der Universität, wie Forschungsgruppen, Institute oder Exzellenz-Cluster einbezogen werden. Die Definition der Rahmenbedingungen an den Standorten ist somit Gemeinschaftsaufgabe von Wissenschaft und Infrastruktureinrichtungen – Ausgangspunkt sollten dabei die Vorgaben und Empfehlungen von Wissenschaftspolitik und Forschungsfördern sein.

Zentrale Entscheidungen, die getroffen und idealerweise in verbindlicher Weise fixiert werden sollten, sind insbesondere:

- Zugang zu Forschungsdaten (intern in der Forschungsgruppe/Einrichtung und extern für Dritte)
- Speicherdauer / Prognoseentscheidung über die zukünftige Bedeutung von Daten und damit ihre Langzeitarchivierung
- Veröffentlichung
- Streitschlichtungsmechanismen
- Festlegung grundlegender Governance-Strukturen und Kompetenzen, insbesondere Klärung der Befugnisse zur Löschung und/oder Zugänglichmachung von Daten

Als Vehikel, um rechtlich bindende Regelungen zu treffen, können dabei Archivierungs-, Veröffentlichungs- und Arbeitsverträge, aber auch Projekt- oder Kooperationsvereinbarungen dienen. Ebenso sind universitäre Satzungen und Verwaltungsvorschriften ein gangbarer Weg. Policies fehlt hingegen eine entsprechende rechtliche Bindungswirkung, gleichwohl sind solche Papiere unerlässlich um den Rahmen des Umgangs mit Forschungsdaten abzustecken.

3.2 Qualitätsmanagement

Empfehlung	Adressat
	Hochschulen
Fachspezifische Regeln zur Datenqualität ausarbeiten und anwenden.	Forschende

Daten formal prüfen, Workflows für die Sicherung der Datenqualität schaffen und z. B. in den Ingest von Repositorien integrieren.	Infrastruktureinrichtungen
-----------------------------------------------------------------------------------------------------------------------------------	----------------------------

Forschungsdaten-Governance und FDM sind nicht Selbstzweck. Vielmehr dienen sie dazu Forschung zu stärken, indem Interdisziplinarität, Überprüfbarkeit und Innovationskraft der Forschung befördert werden. Dafür entscheidend ist aber die Qualität der tatsächlich zirkulierenden Daten. Im Hinblick auf die Prozessierbarkeit und damit intensive Nutzung von Daten sind in diesem Kontext die FAIR-Prinzipien von zentraler Bedeutung.¹¹ Qualitätsmanagement muss aber weitergehen und insbesondere auch die wissenschaftliche Relevanz und inhaltliche Qualität von Daten berücksichtigen. Zu diesem Zweck sind Fach-Communities und Forschende gefordert, eingeübte Qualitätssicherungsmechanismen auf den Kontext digitaler Daten zu übertragen. Service- und Hosting-Einrichtungen können dabei einen wesentlichen Beitrag leisten, indem sie Workflows anbieten, die Forschende und Forschungsgruppen dabei unterstützen Qualitätssicherungsprozesse effizient und nachvollziehbar zu organisieren. Das könnte beispielsweise in Form eines Peer-Reviews im FDM-System umgesetzt sein, wo erst nach einer Freigabe Daten tatsächlich sichtbar werden.

3.3 Daten-Kuratierung

Empfehlung	Adressat
	Hochschulen
Datenkuratierung; insbesondere Vergabe möglichst umfangreicher Metadaten	Forschende
Unterstützung bzw. Mitwirkung an der Datenkuratierung	Infrastruktureinrichtungen

Wie oben angeführt ist die inhaltliche Bewertung der Qualität von Forschungsdaten Sache der Wissenschaft. Sehr wohl aber können Infrastruktureinrichtungen einen Beitrag zur Kuratierung von Daten leisten. Ob dies aktive Mitarbeit oder lediglich die Bereitstellung von Speichersystemen und anderen technischen Strukturen umfasst, hängt vom Einzelfall ab. In jedem Fall sollten Rollen und Zuständigkeiten zwischen Forschenden und Infrastruktureinrichtungen frühzeitig festgelegt werden.

Entscheidend ist aber nicht zuletzt, dass Infrastrukturen darauf hinwirken, dass frühzeitig eine Datenkuratierung erfolgt, denn nach Abschluss von Forschungsprojekten drohen Informationen zu Datensätzen, ihrer Erhebung und der zugrundeliegenden Methodik verloren zu gehen. Daher ist eine gründliche Dokumentation durch Vergabe möglichst umfangreicher Metadaten unverzichtbares Element eines guten FDM.

¹¹ Vgl. Gehring/Lange: Bitte nicht auf Autopilot schalten!, 2.10.2020, <https://www.jmwiarda.de/2020/10/02/bitte-nicht-auf-autopilot-schalten/> [8.10.2020].

3.4 Rechtemanagement

Empfehlung	Adressat
Schaffung von rechtlichen Rahmenbedingungen (z.B. Klauseln in Standardarbeitsverträgen) und Beförderung der Kooperation zwischen den beteiligten Akteuren beim Rechtemanagement.	Hochschulen
Rechtliche Themen bei der Planung von Forschungsprojekten berücksichtigen und auf Lösungen hinwirken.	Forschende
Etablierung von Workflows und Beratungsangeboten für das Rechtemanagement (in Zusammenarbeit mit anderen Stellen wie z.B. der Rechtsabteilung).	Infrastruktureinrichtungen

Zur Sicherstellung und Umsetzung der unter 1.3 genannten rechtlichen Rahmenbedingungen sollte ein aktives Rechtemanagement für Forschungsdaten etabliert werden. Dies umfasst:

- Mechanismen zur Auswahl und Vergabe von Lizenzen
- Arbeits-/dienstrechtliche Regelung des Zugriffs auf Forschungsdaten innerhalb eines Forschungsprojektes
- Verwertung und Sicherung von Erfindungen durch Patentierung

Dabei wirkt erschwerend, dass die zugehörigen Kompetenzen an Universitäten über verschiedene Einrichtungen verstreut sind. Rechtsabteilung, Rechenzentren, Bibliotheken, Forschungssupport und Technologietransfer sind hier aufgefordert zusammenzuwirken, um Rechtssicherheit für die Forschenden zu schaffen und rechtliche Beratung bereitzustellen. Unverzichtbar sind nicht zuletzt fest etablierte Workflows, die dafür sorgen, dass das Rechtemanagement fest in den Prozessen des Datenlebenszyklus verankert und für die Forschenden mit möglichst geringem Aufwand verbunden ist.

Hingegen obliegt es Forschenden, nicht nur bei Projekt- und Förderanträgen, sondern während der gesamten Dauer von Forschungsvorhaben rechtliche Aspekte zu bedenken und auf praxistaugliche Regeln zu hinzuwirken. Begrüßenswert wäre in diesem Kontext, wenn Forschenden in den Selbstverwaltungsgremien der Hochschule auf den Ausbau der vielerorts immer noch unzureichenden Beratungsinfrastruktur hinwirken.

4. Kosten

Empfehlungen	Adressat
Kostenbeteiligungsmodelle sollten Open Science gezielt fördern und gleichzeitig wissenschaftsfreundlich sein, z.B. indem für Langzeitarchivierung Einmalzahlungen vorgesehen werden.	Hochschulen und Infrastruktureinrichtungen

Kosten des FDM schon während der Projektplanung berücksichtigen, in Datenmanagementplänen dokumentieren und in Förderanträgen geltend machen.	Forschende
Die Anwendung des Verursacherprinzips fördert den sparsamen Umgang mit Ressourcen.	Alle

Neben Verantwortung für die Daten gehen mit dem FDM Kosten einher. Die Kosten sollten transparent verteilt und mit einer nachhaltigen Finanzierung unterlegt werden. Eine Refinanzierung kann sich aus verschiedenen Quellen speisen

Produzenten von Forschungsdaten sollten Kosten, die in den verschiedenen Phasen des Datenlebenszyklus anfallen, schon zu Beginn der Projektplanung berücksichtigen. Insbesondere sind die Kosten für Erhebung, Speicherung, Dissemination und Archivierung relevante Faktoren. Wie bereits thematisiert, verändern sich dabei die Rollen der Beteiligten während der Phasen. Diese Veränderungen zeigt sich sowohl in der „Buchhaltung“ sowie der Wertzumessung von Forschungsdaten im Zeitverlauf.

Die grundlegende Gegenüberstellung von Kosten für die Verarbeitung von Daten und dem Wert, der aus dieser Verarbeitung entsteht, ist anwendbar in jeder Phase des Lebenszyklus, aber auch in der Gesamtschau. In anderen Zusammenhängen wird für diese Gesamtschau der Begriff „Total costs of Ownership“ (TCO) angeführt.

Bei der Planung von Forschungsdaten sind einige Kostenarten recht gut abschätzbar. So sind die reinen Speicherkosten, die durch das Schreiben und Bewahren von Bits auf physikalischen Medien entstehen, eine abhängige Größe der Speichermenge.¹² Die Grenzkosten für zusätzliche Speichereinheiten werden mit größeren Mengen tendenziell sinken. Für die Administration der IT-Systeme, zu denen Aspekte wie Betrieb von Serverräumen mit Erfüllung von Sicherheitsanforderungen und die Anstellung von Fachpersonal gehört, werden weitere Kosten entstehen. Die Kosten für den IT-Betrieb können niedriger ausfallen, wenn sie von Einrichtungen mit Fachexpertise und mit einer Mindestgröße gefahren werden, die technisch notwendige Fixkosten besser umlegen können. Zwischen verschiedenen Technologien gibt es Preisunterschiede, die im Zielkonflikt stehen. Schnelle Speicherung und rasche technische Wiederbereitstellung von Daten sind teurer als Systeme mit längeren Zugriffszeiten. Systeme mit verzögerten Zugriffszeiten senken die Kosten für die reine Speicherung, jedoch steigt der Aufwand durch die langsamere und komplexere Wiederbereitstellung. Schnelle Systeme verwenden SSD-Speicher, sehr langsame Systeme verwenden Bandlaufwerke.

Weitere Kosten entstehen durch die Prozesse der Sammlung, Kuratierung, Aufbereitung und Metadatenanreicherung. Sie entstehen in der Domäne der Forschenden während des Vorhabens. Die Kalkulation sollte durch die Forschenden vorgenommen und im DMP dokumentiert werden. Dieser Schritt ist notwendig für einzeln Forschende wie für Gruppen.

¹² Große Datenmengen sind wegen des benötigten Speicherplatzes schneller ein größeres Problem als verhältnismäßig kleine Datensätze. Die könnte man einfach unabhängig von ihrer Popularität mitschleifen, da die Mehrkosten pot. vernachlässigbar.

Bei der Übernahme in weitere Systeme für die Langzeitarchivierung ändert sich die Kostenstruktur. Es ändert sich auch der Einfluss, den Forschende auf den Impact durch ihre Daten haben. Die Kosten für die Langzeitarchivierung wird in aller Regel nicht mehr bei Ihnen selbst liegen. Sie werden durch die Hochschule oder durch andere Einrichtungen übernommen, die die Langzeitarchivierung aus eigener Priorität heraus oder im Auftrag organisieren. Hochschulen werden den Wert in der Relevanz der Daten oder in der Frequenz der Nachnutzung sehen. Wenn die Kosten der Archivierung durch gute Vorstrukturierung und wenig komplexe Lizenzen reduziert werden, wird auch bei niedriger bewerteten Daten eine Speicherung erwogen werden können.

Die Stellschraube, mit der Forschende eine langfristige Speicherung auch nach Verlassen der Daten aus ihrer Sphäre wahrscheinlicher machen können, ist die vereinfachte Nachnutzung durch gute Strukturierung, einfache Nachvollziehbarkeit (durch Metadaten) und zugängliche Lizenzmodelle.

Die Anwendung des Verursacherprinzips ist insbesondere bei großen Datenmengen in gewissem Umfang sinnvoll, um einen verantwortlichen Umgang mit den Ressourcen sicherzustellen. Dieses wurde beispielsweise mit bwSFS (Storage for Science) an den Standorten Freiburg und Tübingen für den gesamten Datenlebenszyklus in einem Großgeräteantrag realisiert. Für zukünftige Anforderungen sollten die Mittel für Langzeitarchivierung genauso wie eventuelle Mittel in der Phase der Datenerhebung und Verarbeitung im Projektantrag ausgewiesen werden. Hierbei sollte eine direkte Übertragung als Erweiterung bestehender zentraler Ressourcen (ob lokal, Community oder NFDI) angedacht werden.

Kostenbeteiligungsmodelle sollten zudem weitestgehend Open-Science-Praktiken unterstützen. Aus Sicht der Forschenden sollte die Publikation von Daten nicht teurer sein als das bloße Archivieren in einem Dark Archive. Dies steht in einem Spannungsverhältnis dazu, dass ersteres in der Regel mit höheren Kosten verbunden ist (publizierte Daten sollen langfristig, ggf. sogar für immer erhalten werden, sie werden u.U. eher auf Festplatte als auf Band gespeichert, um performant verfügbar zu sein, sie erfordern einen höheren Aufwand an Aufbereitung und Kuration). Hier sind die Institutionen gefordert, entsprechende Modelle über die Budgets ihrer Infrastruktureinrichtungen zu ermöglichen. Umgekehrt sollten die Infrastruktureinrichtungen möglichst Kostenbeteiligungsmodelle anstreben, die Einmalzahlungen von Seiten der Forschenden vorsehen, die sich dann auch beispielsweise im Rahmen von Projektanträgen einwerben lassen. Eine zusätzliche Herausforderung erwächst aus dynamischen oder interaktiven Komponenten eines Projekts. So werden Regelungen zum Umgang mit entwickelten Tools und Services (Web-Schnittstellen, Interfaces) benötigt und Punkte wie Portierungsmöglichkeiten, Interaktivität, langfristige Verfügbarkeit trotz Wegfall der bestehenden Projektfinanzierung sind zu berücksichtigen.

Ausblick

Das Thema Forschungsdaten hat in den vergangenen Jahren stetig an Dynamik gewonnen und wird im Hochschulalltag inzwischen in Breite wahrgenommen. Dennoch ist der Prozess der Etablierung von Standards und Best-Practices längst noch nicht abgeschlossen. Dieser Leitfaden versucht, mit aktuellen praktischen Empfehlungen den Akteuren Handlungsoptionen aufzuzeigen und gleichzeitig einen Beitrag zur Diskussion um eine Standardisierung von FDM-Prozessen und -Strukturen auf den verschiedenen Ebenen vorzubringen.



Insbesondere gilt es, für Großprojekte wie NFDI oder EOSC auf lokaler Ebene Grundlagen zu schaffen. Voraussetzung für den Erfolg solcher kooperativen, institutionsübergreifenden Netzwerke ist, dass FDM-Standards in Forschungsprojekten und an Hochschulen etabliert und eingeübt sind. Dies gilt umso mehr als sich Projekte wie die NFDI ihrerseits nach außen öffnen und internationale Anschlussfähigkeit anstreben. Angesichts einer international agierenden und kooperierenden Forschung ist dies eine Notwendigkeit. Damit werden lokale FDM-Maßnahmen und -strukturen nicht überflüssig. Im Gegenteil: einrichtungsübergreifende und lokale Strukturen komplementieren und benötigen sich gegenseitig.