

## Handlungsempfehlungen zum Forschungsdatenmanagement an der Universität Erfurt *(Stand: 18. Dezember 2019)*

### Vorbemerkungen

Die Universität Erfurt fördert den freien Zugang zu Forschungsdaten und ist als Mitglied der Deutschen Forschungsgemeinschaft (DFG) den im Kodex 2019 ausformulierten „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ verpflichtet. Der Senat der Universität Erfurt hat am 03. Juli 2019 einstimmig eine [Forschungsdatenmanagement-Policy](#) für die Universität verabschiedet, in welcher die Grundsätze im Umgang mit Forschungsdaten beschrieben sind. Die Handlungsempfehlungen ergänzen diese Grundsätze und geben praktische Hinweise zur Umsetzung.

Um einen transparenten, verantwortungsvollen und wissenschaftlich angemessenen Umgang mit Forschungsdaten zu gewährleisten, hat die Universität Erfurt eine zentrale Unterstützungsstruktur in Form einer Servicestelle „Forschungsdatenmanagement“ etabliert, die Mitgliedern und Angehörigen der Universität als Ansprechpartner für alle Fragen zum Thema „Forschungsdaten“ zur Verfügung steht. Bei der Gestaltung der Serviceangebote orientiert sich die Servicestelle an nationalen und internationalen Entwicklungen und kooperiert im Rahmen des Thüringer Kompetenznetzwerks Forschungsdatenmanagement mit den anderen Hochschulen in Thüringen.

Als Forschungsdaten im Sinne dieser Handlungsempfehlungen werden alle Daten bezeichnet, die Gegenstand eines Forschungsprozesses sind, während eines Forschungsprozesses entstehen oder sein Ergebnis sind. Sie werden abhängig von der Forschungsfrage und unter Anwendung verschiedener Methoden in unterschiedlichen Medientypen, Aggregationsstufen und Formaten erzeugt bzw. gesammelt, gespeichert, bearbeitet, analysiert und schließlich publiziert und/oder archiviert. Den FAIR-Prinzipien folgend, sollten Daten **F**indable, **A**ccessible, **I**nteroperable und **R**e-Usable sein.

## I. Planungsphase: vor dem Forschungsvorhaben

- Wenn Sie ein Forschungsvorhaben planen, ist eine frühzeitige Befassung mit einem effizienten und nachhaltigen Forschungsdatenmanagement (FDM) dringend angeraten. Nutzen Sie das Beratungsangebot der Servicestelle Forschungsdatenmanagement, um eine geeignete Strategie für Ihr Vorhaben zu entwickeln und Ihren individuellen Bedarf zu ermitteln, etwa zu Fragen eines FDM-bezogenen Supports bei Fördermitteln/Projekt-Anträgen, der Erstellung eines Datenmanagementplans, der Publikation Ihrer Forschungsdaten oder der dauerhaften Sicherung und Archivierung Ihrer Forschungsdaten.
- Empfohlen, und bei einigen Förderorganisationen auch gefordert, ist die Erstellung und Einreichung eines Datenmanagementplans (DMP).<sup>1</sup> Er beschreibt die Anforderungen sowie die technischen und organisatorischen Maßnahmen, die während der Projektlaufzeit, aber auch nach Abschluss des Projektes den nachhaltigen Umgang mit Forschungsdaten gewährleisten sollen. Er stellt ein zentrales Dokument der von der DFG geforderten phasenübergreifenden Qualitätssicherung dar. Ein DMP berücksichtigt auch datenschutzrechtliche Aspekte und dokumentiert so den Umgang mit den Anforderungen der Datenschutz-Grundverordnung (DSGVO).
- Folgende Fragen können hilfreiche Anhaltspunkte für die Erstellung eines projektspezifischen Datenmanagementplans bieten:
  - Was ist das Ziel des Projekts? Welche Einrichtungen / Personen werden miteinbezogen?
  - Welche Forschungsdaten werden dafür erhoben / nachgenutzt und wie?
  - Welche (fachspezifischen) Standards werden angewendet / berücksichtigt (z.B. Datenformate, Metadaten, Ontologien)?
  - Wie wird die Speicherung, Sicherung, Archivierung und ggf. Zugänglichmachung von Forschungsdaten organisiert und dokumentiert?
  - Wie wird getestet, ob die Erfordernisse erfüllt werden, in welchem zeitlichen Rhythmus finden diese Tests statt, und wie wird auf sie reagiert?
  - Gibt es rechtliche, zeitliche oder andere Einschränkungen für die Zugänglichmachung von Forschungsdaten, die berücksichtigt werden müssen?
- Forschungsförderorganisationen erwarten in der Regel nicht nur die langfristige Sicherung von Forschungsdaten, sondern auch deren Veröffentlichung. Eine geeignete Strategie sollte daher bereits in der Antragsphase bestimmt werden. Die dabei entstehenden Kosten (z. B. für Personal, Speicherung, Publikation), die über die Grundausstattung der Universität hinausgehen, können und sollten Teil der beantragten Mittel sein. Wenden Sie sich bitte an den Stabsbereich Forschung

---

<sup>1</sup> Macht die avisierte Förderrichtlinie Vorgaben zu Struktur und Details des Datenmanagements, sind diese unbedingt zu berücksichtigen.

und Nachwuchsförderung, wenn im Rahmen der Antragstellung die Möglichkeit der Beantragung solcher Mittel nicht vorgesehen ist. Bei der Planung und Beantragung von ggf. notwendigen Geräten und IT-Dienstleistungen ist das Universitätsrechen- und Medienzentrum (URMZ) schon in der Vorbereitungsphase von Anträgen einzubeziehen.

- Bei vielen Forschungsvorhaben und insbesondere bei der Veröffentlichung der Ergebnisse sind rechtliche Rahmenbedingungen zu beachten. Für Studien an Personen bzw. Personengruppen ist ggf. das Votum des Ethikbeirats einzuholen, um Rechte und Sicherheit der Probanden im Sinne der Deklaration von Helsinki zu schützen.
- Personenbezogene Daten, insbesondere sensible Daten sind nach den Vorgaben der DSGVO zu verarbeiten. Auch der Schutz des Urheberrechts und die berechtigten Interessen Dritter müssen gewährleistet sein. Deshalb sind bei der Planung eines Forschungsvorhabens grundlegende rechtliche Fragen im Voraus abzuklären.
- Die Nutzung innerhalb eines Forschungsprojektes generierter Daten steht den erhebenden Forschenden zu: Bei Forschungsvorhaben, die mehrere Kooperationspartner betreffen, sind Vereinbarungen über die Nutzungsrechte in der Antragsphase zu treffen. Wenn der die Daten erhebende Forschende die Universität Erfurt verlässt, ist der weitere Umgang mit den Nutzungsrechten vertraglich zu klären.

## II. Durchführungsphase: während des Forschungsvorhabens

- Der in der Antragsphase erstellte DMP dient in der aktiven Projektphase als Referenz und ist ggf. an veränderte Rahmenbedingungen und Anforderungen anzupassen (Stichwort: living document). Insbesondere sind die Rollen und Verantwortlichkeiten der am Projekt Beteiligten fortlaufend zu dokumentieren, um wechselnde Verantwortlichkeiten nachvollziehen zu können.
- Die Dokumentation der Forschungsergebnisse umfasst alle für ihr Zustandekommen relevanten Informationen, soweit dies zumutbar ist, um Arbeitsläufe umfänglich darzulegen. Dies schließt zugrundeliegende Forschungsdaten, Materialien und Informationen, die angewandten Methoden und Arbeitsabläufe sowie selbst programmierte Software und Informationen zu genutzten kommerziellen Softwareprodukten explizit ein. Die Dokumentation umfasst grundsätzlich alle Ergebnisse – auch jene, die die Forschungshypothese nicht stützen – und ist bestmöglich gegen Manipulationen zu schützen.

- Für die Verarbeitung und Speicherung von Forschungsdaten sowie die Kollaboration auf Basis dieser Daten sollten Verfahren auf dem Stand der Technik verwendet werden. Dazu gehört insbesondere die Beachtung von Datensicherheit hinsichtlich Verfügbarkeit, Integrität (Unverfälschtheit) und Authentizität. Hierzu gehören z. B. die Nutzung von Datensicherung und Archivierung, die Verwendung sicherer Datenaustauschplattformen und eine nachvollziehbare Dokumentation des Versionsverlaufs z. B. über die Verwendung von Versionierungstools.
- Wichtig ist, die Effektivität der eigenen Datensicherungsstrategie und die Performance der verschiedenen Kollaborationsstrategien während des Projektes kontinuierlich zu testen und gegebenenfalls die Strategien anzupassen.
- Das URMZ stellt im Forschungsdatenumfeld u. a. nachfolgende Dienste zur Verfügung. Die Dienste werden teilweise als Grundausstattung kostenfrei angeboten oder können zu Grenzkosten zur Verfügung gestellt werden.
  - Nutzung von Netzwerkfilesystemen (inkl. Datensicherung)
  - Bereitstellung von virtuellen Servern
  - Datenaustauschdienste (DFN-Cloud)
  - Langzeitarchivierung in Kooperation mit der Thüringer Universitäts- und Landesbibliothek (ThULB)

Nähere Dienstbeschreibungen finden Sie auf den Webseiten des URMZ.

- Von elementarer Bedeutung für die Nachnutzbarkeit von Forschungsdaten ist deren Beschreibung mit Metadaten. Metadaten sind Daten über Daten und geben den Kontext wieder, in dem diese erzeugt wurden. Als Faustregel gilt: Metadaten sollten die klassischen sechs W-Fragen beantworten:
  - Wer?
  - Was?
  - Warum?
  - Wie?
  - Wann?
  - Wo?

Metadaten sind die Voraussetzung, um potenziellen Nachnutzern das Auffinden und die Abschätzung der Eignung für die Nutzung zu ermöglichen. Idealerweise erfolgt die Beschreibung strukturiert und maschinenlesbar. Hierzu existieren in den meisten Fachdisziplinen Metadatenstandards und standardisierte Terminologien. Falls diese nicht existieren, sollten allgemeingültige Standards, wie z. B. Dublin Core, MARC oder MODS zur Beschreibung verwendet werden.

- In der Durchführungsphase eines Projektes unterliegen Datensätze meist einer Entwicklung über mehrere Stufen hinweg (z. B. durch Auswahl, Aggregation, Integration). Es hat sich als vorteilhaft erwiesen, die verschiedenen Versionen zu

kennzeichnen, zu dokumentieren und zumindest während der aktiven Projektlaufzeit aufzubewahren. Insbesondere bei textbasierten Daten erleichtert die Verwendung von Versionierungstools, wie sie in der Softwareentwicklung üblich sind (z. B. git, SVN), das Management der verschiedenen Versionen.

- In Verbundprojekten oder bei großen Datenmengen ist der Einsatz dezidierter Arbeitsumgebungen und Portale zum Datenmanagement sinnvoll. Der Betrieb dieser Tools erfordert meist zusätzliche Ressourcen, bietet jedoch den Vorteil einer einheitlichen und zentralen Verwaltung der Forschungsdaten. Das Auffinden und der Austausch von Daten wird damit erleichtert, sollte jedoch mit Hilfe eines abgestimmten Datenmanagementschemas zu bspw. Dateibenennung, Kodierung, etc. geregelt werden.

### III. Abschlussphase: nach dem Forschungsvorhaben

- Während die Dokumentation eines Projektes den gesamten Projektumfang erfasst, treffen Forschende bei der Veröffentlichung eine Selektionsentscheidung. Sie entscheiden in eigener Verantwortung – unter Berücksichtigung der Standards des jeweiligen Fachgebiets –, ob, wie und wo sie ihre Ergebnisse öffentlich zugänglich machen und legen diese Entscheidungen nachvollziehbar dar.
- Im Sinne der guten wissenschaftlichen Praxis sind Forschungsdaten spätestens zum Projektende langfristig zu sichern und zugänglich zu machen. Insbesondere auf die Zugänglichkeit wird besonderer Wert gelegt, um die Nachvollziehbarkeit von Forschungsergebnissen sowie eine Nachnutzung der Daten zu ermöglichen. Es wird empfohlen, Forschungsdaten nach dem Prinzip: „Zugänglich, wenn möglich, eingeschränkt, wenn nötig“ zu veröffentlichen.
- Für die Veröffentlichung von Forschungsdaten stehen Ihnen verschiedene Wege zur Verfügung. Neben allgemeinen fachübergreifenden Repositorien existiert eine große Anzahl fachspezifischer Repositorien. Das Portal [re3data.org](https://re3data.org) bietet einen guten Überblick mit umfangreichen Such- und Filterfunktionen.
- Als Mitglieder oder Angehörige der Universität Erfurt können Sie Forschungsdaten über die Digitale Bibliothek Thüringen (DBT) veröffentlichen und archivieren. Es wird empfohlen, alle externen Datenpublikationen auch in der DBT zu registrieren.
- Publikationen sollten als Mindeststandard den FAIR Kriterien entsprechen. Darüber hinaus haben Forschende Freiheiten bei der Wahl der vergebenen Lizenzen und Nutzungsrechte der Daten, der Bekanntheit und Sichtbarkeit der

gewählten Publikationsart innerhalb ihrer relevanten Disziplin sowie den Kosten, die ihnen durch eine Veröffentlichung entstehen.

- Forschungsdaten sollten in einer Verarbeitungsstufe (Rohdaten oder bereits weiter strukturierte Daten) zugänglich sein, die eine sinnvolle Nach- und Weiternutzung durch Dritte ermöglicht. Insbesondere Daten, die die Grundlage wissenschaftlicher Artikel bilden, sollten zugänglich gemacht werden.
- Wie bei wissenschaftlichen Artikeln sollten auch Forschungsdaten mit der Veröffentlichung einen eindeutigen, persistenten Identifikator (PID) zugewiesen bekommen. Die Vergabe einer PID (z. B. DOI, URN) erfolgt in der Regel durch das Repositorium oder das Data Journal, in dem die Daten publiziert werden.
- Um die Nutzungs- und Verwertungsrechte von Forschungsdaten zu regeln, sollten Daten immer mit einer entsprechenden Lizenz veröffentlicht werden. Die Wahl der Lizenz sollte einen offenen Zugang für wissenschaftliche Zwecke erlauben. Zu berücksichtigen sind ggf. die speziellen Anforderungen der Förderorganisation bzw. Repositorien. Geeignet sind die Open Data Commons bzw. die Creative Commons Lizenzen sowie die GNU General Public License (GPL) für Software.
- Die Veröffentlichung von Forschungsdaten sollte so zeitnah wie möglich erfolgen. Liegen triftige Gründe vor, können Daten mit einer Sperrfrist veröffentlicht werden. Die Festlegung der Sperrfrist erfolgt durch die verantwortlichen Forschenden. Es gilt jedoch, die Vorgaben und Richtlinien der Forschungsförderorganisationen und Repositorien zu beachten. Sperrfristen sollten eine Dauer von 1-2 Jahren nach Projektende nicht übersteigen.
- Der Archivierung von Forschungsdaten liegen die Empfehlungen der DFG zu einer Aufbewahrung von mindestens 10 Jahren nach Projektende zugrunde. Kürzere Aufbewahrungsfristen für Teile der Daten sind im DMP zu erläutern, der gemeinsam mit den Daten(UMRZ) abzulegen ist.

## Kontakt

Dr. Nadine Neute  
Servicestelle Forschungsdatenmanagement der Universität Erfurt  
Thüringer Kompetenznetzwerk Forschungsdatenmanagement (TKFDM)

Universitätsbibliothek, Raum 140  
Nordhäuser Straße 63  
99089 Erfurt

Tel.: +49 (0)361 737-5707

E-Mail: [forschungsdatenmanagement@uni-erfurt.de](mailto:forschungsdatenmanagement@uni-erfurt.de)